Characterisation of human Haematopoietic Stem Cell differentiation and exit from quiescence at single cell resolution

Serena Belluschi

Darwin College

University of Cambridge

December 2019

This dissertation is submitted for the degree of Doctor of Philosophy

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any work that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Abstract

Serena Belluschi

Characterisation of human Haematopoietic Stem Cell differentiation and exit from quiescence at single cell resolution

Blood formation is coordinated by a set of functionally heterogeneous haematopoietic stem cell (HSC) compartments both in mouse and in human. However, how this diversity is regulated at the cellular and molecular level and when HSC multipotency is lost is still not known, in particular in humans. Quiescence is an important property of HSCs and fine regulation of the balance between quiescence and cell cycle entry is of vital importance for maintaining a healthy HSC pool and avoid haematological malignancies. Very little is known about how molecular networks change during exit from quiescence and no studies have formally examined if cell fate decisions occur during this process or later on during the cell cycle.

Here, I combine index sorting, *in-vitro* single cell functional assays and single cell RNA-sequencing with xenotransplantation assays to profile single human HSC properties in the purest HSC compartment reported to date ($CD49^+$ HSCs) and to understand when the first steps of lineage restriction occur during HSC differentiation. Moreover, I use an *in-vitro* model system to comprehensively study HSC activation and to investigate whether HSC self-renewal is lost during quiescence exit or during cell cycle progression.

First, I unveil an unexpected degree of intrinsic functional and molecular heterogeneity within the human $CD49f^+$ HSC pool. I demonstrate that the first restriction step towards the lymphoid lineage occurs already within the $CD49f^+$ HSC compartment and generates erythroid-deficient myeloid-lymphoid committed cells. Within this compartment, transcriptional programmes and lineage potential progressively change along a gradient of opposing cell surface expression of CLEC9A and CD34. Two functionally distinct populations can be identified and purified. $CLEC9A^{hi}$ $CD34^{lo}$ cells contain long-term repopulating multipotent HSCs with slow quiescence exit kinetics, whereas $CLEC9A^{lo}$ $CD34^{hi}$ cells are restricted to myelo-lymphoid differentiation and display infrequent but durable repopulation capacity.

Second, I demonstrate that a drastic transcriptional remodelling reflective of the fast metabolic activation seen in *ex-vivo* cultured HSCs occurs during quiescence exit, and independently of cell cycle progression. *In-vivo* data show that the reduction in repopulation capacity that accompanies HSC culture also occurs independently of cell cycle progression. Recent work highlighted the active role of mitochondria and metabolism in HSC fate decision and self-renewal. My data is consistent with a model in which the metabolic remodelling seen during exit from quiescence represents the first step towards loss of self-renewal and differentiation. This has important implications for improving *ex-vivo* protocols for HSC culture and HSC transplants.

Acknowledgments

Firstly, I would like to express my gratitude to my supervisor Dr. Elisa Laurenti for being a great mentor and for giving me guidance and support through the past four years. It has been an amazing experience and a great journey, from being the first PhD student in the lab to seeing the lab expanding and accomplishing great science. Thanks for transmitting me her passion for science and for sharing her expertise, for being patient and always positive and supportive, also in the more difficult times. I have grown as a person and as a scientist and I feel ready to take the next step in my career.

I would very much like to thank my second supervisor, Professor Bertie Göttgens for his invaluable advice and support.

Thanks to all the members of the Laurenti Lab, to the ones who left and the current ones, for being great colleagues and friends, for your help and advice and for making such a great team. A special thanks to Emily Calderbank for being a great lab manager and for working along with me up to the publication of our co-authored manuscript. Thanks also for delighting us with your amazing cakes! A special thanks goes also to Carys Johnson for helping with the project and for continuing part of the work presented here. Thanks to Michael Drakopulous for the experiments and research you have performed for part of the work presented here.

My sincere thanks go to the present and past members of the Gottengs Lab for sharing their expertise on single cell RNA sequencing. Thanks to Winnie Lau, Fiona Hamey, Nicola Wilson and Fernando Calero. Thanks to Kenny Sham, Evangelia Diamanti and Xiaonan Wang for your fantastic work and analysis of RNA sequencing data, without which most of this work wouldn’t have been possible.

I also want to thank Miriam and Antonella, colleagues that became best-friends. Thanks for always being there for me!

Finally, I want to thank my family that despite the distance is always with me. A special thanks to Ivan, for his constant encouragement, love and support and to my little daughter, for spreading her joy and happiness and giving me the strength to do more every day.

Table of contents

List of abbreviations

BM	Bone Marrow
°C	Celsius
2-NBDG	2-Deoxy-2-[(7-Nitro-2,1,3-Benzoxadiazol-4-Yl) Amino]-Dglucose
a-HSC	Activated-HSC
Bala	Balanced
BF	Bright Field
BrdU	Bromodeoxyuridine
BSA	Bovine Serum Albumin
Ca^{2+}	Calcium
CB	Cord Blood
CBSB	Cambridge Blood and Stem Cell Biobank
CDK	Cyclin-Dependent Kinases
CFU-C	Colony Forming Unit-Cell
CFU-S	Colony Forming Unit-Spleen
CKI	Cyclin-Dependent Kinases Inhibitor
CLP	Common Lymphoid Progenitors
CMP	Common Myeloid Progenitor
CP	Canonical Pathways
CRU	Competitive Repopulating Unit
CV	Coefficient of Variance
DC	Dendritic Cell
DMEM	Dulbecco's Modified Eagle Medium
ELDA	Extreme Limiting Dilution Analysis
Ery	Erythroid/Erythrocyte
ESC	Embryonic Stem Cells
ETP	Early T-Cell Precursor
FACS	Fluorescence-Activated Cell Sorting
FBS	Fetal Bovine Serum
FL	Fetal Liver
FUCCI	Fluorescence Ubiquitin Cell Cycle Indicator
G_1	Gap1 Phase
G_2	Gap2 Phase
GMP	Granulocyte-Monocyte Progenitor
GO	Gene Ontology

Gran	Granulocytic/Granulocyte
GSEA	Gene Set Enrichment Analysis
hEPO	Human Erythropoietin
HIF-1α	Hypoxia Inducible Factor-1 Alpha
HSC	Haematopoietic Stem Cell
HSCT	Haematopoietic Stem Cell Transplants
HSPC	Haematopoietic Stem and Progenitor Cell Compartment
HTSeq	High-Throughput Sequencing
HVGs	Highly Variable Genes
IB	Injected Bone
ICGS	Iterative Clustering and Guide-Gene Selection
IF	Intrafemorally
IFN	Interferon
IL-1R	Interleukin-1 Receptor
IMDM	Iscove's Modified Dulbecco's Medium
IT-HSC	Intermediate-Term Haematopoietic Stem Cell
LDA	Limiting Dilution Assay
Lin	Lineage
LMPP	Lymphoid-Primed Multipotent Progenitor
lncRNA	Long Non Coding RNA
LSK	Lineage- Sca-1^+ C-Kit^+
LT-HSC	Long-Term Haematopoietic Stem Cell
LT-IC	Long-Term Initiating Culture
Ly	Lymphoid/Lymphocyte
Ly-bi	Lymphoid-Biased
M	Mitotic Phase
Meg	Megakaryocytic/Megakaryocyte
MEM	Myeloid-Erythroid-Megakaryocytic
MEP	Megakaryocyte-Erythrocyte Progenitor
MFN2	Mitofusin 2
MLP	Multilymphoid Progenitor
MNC	Mononuclear Cell
Mon	Monocytic/Monocyte
MonGran	Monocytic-Granulocytic/Monocyte-Granulocyte
MPP	Multipotent Progenitor
MuSCs	Muscle Stem Cells

My	Myeloid
My-bi	Myeloid-Biased
NR	Nicotinamide Riboside
NSC	Neural Stem Cells
NSG	Nod Scid Gamma
pRB	Phosphorylated Retinoblastoma
PB	Peripheral Blood
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PD	PD0332991, Palbociclib
pq-HSC	Pharmacologically Quiescent HSC
QC	Quality Control
q-HSC	Quiescent-HSC
RA	Retinoic Acid
RB	Retinoblastoma
RBC	Red Blood Cell
RPL	Ribosomal Protein L
RPS	Ribosomal Protein S
RT	Room Temperature
S	Synthesis Phase
Sca-1	Stem Cell Antigen-1
SCF	Stem Cell Factor
scRNA-seq	Single Cell RNA-Sequencing
SP	Side Population
SPF	Specific-Pathogen-Free
SRC	Scid Repopulating Cell
STEM	Short Time-Series Expression Miner
ST-HSC	Short-Term Haematopoietic Stem Cell
TF	Transcription Factor
TFT	Transcription Factor Targets
TGF-β	Transforming Growth Factor B
TMRM	Tetramethyl rhodamine Methyl Ester
TPO	Thrombopoietin
tSNE	T-Distributed Stochastic Neighbour Embedding
UMAP	Uniform Manifold Approximation and Projection

1 INTRODUCTION

1.1 Haematopoiesis and haematopoietic stem cells

The production of all mature blood cell types during an individual lifetime is constantly maintained by a process called haematopoiesis. Haematopoiesis in adults occurs primarily in the bone marrow (BM) and it produces approximately one trillion cells every day (Doulatov et al. 2012). Haematopoiesis is one of the best-described models for the study of stem cells due to the ease of blood collection and cell isolation.

The advent of flow cytometry and the development of clonal *in-vivo* and *in-vitro* functional assays contributed to a better understanding of haematopoiesis, which is generally described as a hierarchical organisation of multipotent and unipotent cells with a common identified precursor residing at the root of the hierarchy, the haematopoietic stem cell (HSC) (Doulatov et al. 2012; Eaves 2015). HSCs are defined as multipotent cells with the ability to self-renew. They can generate more committed daughter cells but also equally potent cells with the same properties of the parental HSC. The ability to self-renew is better defined by the capacity of serial long-term reconstitution of all blood lineages upon transplantation into a myeloablated recipient. The ability to generate daughter stem cells is at the basis of the differentiation process that leads to a gradual cell fate restriction through the production of progenitors cells and differentiated mature blood cells (Eaves 2015). Another important property of HSC is the ability to reside outside of the cell cycle in a state called quiescence. Quiescent HSCs divide only once every 145 days in homeostatic conditions in mice whereas it has been estimated that they divide every 280 days in the human BM (A. Wilson et al. 2008; Catlin et al. 2011). Human HSCs do not only reside in the BM but can also be isolated from cord blood (CB) or peripheral blood (PB). They are very rare cells and their frequency varies between tissues in human, with CB having the highest frequency (1 in $9.3x10^5$ cells), followed by BM (1 in $3x10^6$ cells) and PB (1 in $6x10^6$) (J. C. Wang, Doedens, and Dick 1997). In mice, HSC frequency has been confirmed by functional studies and is estimated to be of 1 in 10^5 BM cells (Szilvassy et al. 1990).

It is of vital importance for HSCs to finely regulate the balance between proliferation and quiescence in order to avoid the exhaustion of their pool and the development of haematological malignancies.

Due to the ability of HSCs to reconstitute the entire blood system upon transplantation, HSC transplants (HSCTs) are nowadays routinely used to treat leukaemia and other haematological malignancies. Despite their use in clinical transplantation, most of our understanding of HSCs is still derived from mouse models and the mouse HSC is probably the best characterised adult stem cell. However, after the first engraftment of primary human HSCs in immune-deficient mice over 20 years ago, xenotransplantation has been routinely used for the study of human HSCs. Xenotransplantation assays have improved over the years thanks to the development of different new humanised mouse models (Kamel-Reid and Dick 1988; Doulatov et al. 2012). Although the use of mouse models has been

indispensable for the study of haematopoiesis, these models are still far from modelling what really happens in humans. The discovery of species-specific differences in HSC biology highlighted the necessity to study human HSCs directly.

The extreme rarity of human HSCs has put some obstacles in the study of these cells. However, the availability of CB tissues, highly enriched HSC sources, and the advances in HSC isolation and purification facilitated the study of this rare population. This makes CB the most studied source of HSCs in human. Other than differences in HSC frequency, differences in composition and properties also exists in HSCs isolated from CB, BM and PB. CB derived cells seems to have greater proliferative response when cultured *in-vitro* in the presence of cytokines and they are less dependent on stromal cells, often used to sustain HSC proliferation and differentiation. A higher number of colony forming cells with high proliferative potential and more primitive HSCs are present in CB compared to BM and PB (Hordyjewska, Popiołek, and Horecka 2015). These studies could reveal even more important difference in their properties than what is already known.

1.1.1 Early evidence and isolation of adult mouse and human HSCs

In the early 1950s, research on the consequences of myeloablation showed that intravenous transplant of mouse BM cells could protect the syngeneic recipient from a lethal dose of radiation regenerating the previously ablated haematopoietic system in the host. This was the first evidence of the existence of HSCs in mouse (Jacobson et al. 1951; Ford et al. 1956). However, the confirmation of the presence of single multipotent cells capable of regenerating any mature blood cells only came later thanks to the seminal work of Till and McCulloch that confirmed the presence of macroscopic colonies of multiple lineages in the spleen of mice transplanted with donor-derived BM cells. These colonies, named colony forming unit-spleen (CFU-S), were derived from a few or even a single progenitor cell capable of forming multiple lineages and they also contained daughter cells with similar characteristics as the parental cell (A. J. Becker, McCulloch, and Till 1963; Till, McCulloch, and Siminovitch 1964). These results introduced for the first time the concept of multipotentiality and self-renewal that will later be definitive characteristics of HSCs. To complement these studies, new short-term *in-vitro* clonogenic assays were developed to assess haematopoietic cells for their function and differentiation capacity. These studies provided first evidence for intermediate differentiation stages between HSCs (measured as CFU-S) and mature blood cell (Pluznik and Sachs 2005; Eaves 2015).

HSC isolation and characterisation were only possible thanks to the availability of new strategies allowing for better separation of these cells from other BM cellular components and the emergence of *in-vivo* competitive repopulating unit assays (CRU) designed to detect single cells capable of reconstituting blood cell production for long-term (Miller, Dykstra, and Eaves 2008). Resistance to irradiation and 5-florouracil, together with Hoechst 33342 dye exclusion were the properties used to isolate HSCs for the first time (Goodell et al. 1996). A better understanding of surface antigen markers and the advent of flow cytometry later allowed for prospective isolation of multipotent cells capable of

reconstituting multiple lineages. These populations, enriched in HSCs, were further characterised based on their lineages reconstitution upon transplantation and assessed for self-renewing properties through serial transplantation (Spangrude, Heimfeld, and Weissman 1988; Ng and Alexander 2017).

Over the past few decades continuous reporting of new surface markers for mouse HSC isolation has made even more difficult to know which isolation strategy can yield the purest fraction of true long-term repopulating HSCs (Challen et al. 2009). It is important to note that despite the type of strategy used for HSC isolation, identification of true HSCs should be proved as gold-standard by *in-vivo* long-term (>16 weeks) repopulating and serial transplantation assays.

Investigation of human HSCs was firstly conducted using an *in-vitro* colony forming unit-cell (CFU-C) assay (Moore, Williams, and Metcalf 1973). Improved culturing conditions and the use of feeder layers cells allowed for the development of long-term initiating culture (LT-IC) assays that were aimed at detecting primitive hematopoietic stem cells based on their capacity to produce myeloid progeny for at least 4 weeks (Sutherland et al. 1991; Miller, Dykstra, and Eaves 2008). The study and isolation of human HSCs, as compared to mouse HSCs, was challenged by the extreme rarity of these cells and the lack of efficient *in-vivo* model systems. Over the past 20 years, improved xenograft models (NSG mice) which supported higher levels of human engraftment allowed to better characterise and isolate human HSCs (M. Ito et al. 2002; Doulatov et al. 2012). Xenograft models are now extensively used to track the lineage output and the self-renewal properties of single HSCs over more than 8 months (Notta et al. 2011).

It is now known that human HSCs diverge in terms of cell surface markers expression when compared to mouse HSCs. While CD34 expression was first identified to enrich for human HSCs and progenitors (Civin et al. 1984), the expression of this marker is minimal in mouse long-term repopulating cells. Further studies confirmed the lack of congruence for most of the expressed cell surface markers between mice and humans (Sitnicka et al. 2003; Larochelle et al. 2011; Doulatov et al. 2012).

All human HSCs are negative (Lin-) for the expression of mature lineage markers (CD2, CD3, CD14, CD19, CD56, GlyA), similarly to the mouse model. However, most of the markers used for HSCs isolation in mouse are not useful in human. Human HSCs have been initially defined by being $CD34^{+}$, $CD38^{-}$ and $CD45RA^{-}$. Although the $CD34^{+}$ haematopoietic stem and progenitor cell compartment (HSPC) compartment is the most studied and characterised to date, several studies questioned whether the $CD34^{-}$ compartment would also have a role in haematopoiesis. A $CD34^{-}$ population with severe combined immunodeficiency repopulation capacity was identified by Dick et al. in 1998 and confirmed by other studies (Bhatia et al. 1997; Danet et al. 2003; J. Wang et al. 2003). More recent studies identified and characterised an immature ad highly quiescent Lin^{-} $CD34^{-}$ $CD38^{-}$ $CD93^{hi}$ population, with self-renewal capacity and similar to previously identified $CD34^{+}$ HSCs but placed above those in the haematopoietic hierarchy (Anjos-Afonso et al. 2013). CD90 and CD49f expression has also been associated with HSC enrichment, even though heterogeneity in their expression has been identified in the HSC pool (Laurenti and Dick 2012). The more differentiated progenitor populations are instead

identified by the expression of CD34 and gradual acquisition of CD38 and CD45RA expression during the differentiation process (Lansdorp, Sutherland, and Eaves 1990; Bhatia et al. 1997; Doulatov et al. 2012). Therefore, the combined expression of Lin^- $CD34^+$ $CD38^-$ $CD45RA^-$ $CD90^+$ $CD49f^+$ ($CD49f^+$ HSCs) is the best purification strategy to date for the isolation of human HSCs. However, only 10% of this population is capable of long-term engraftment in xenograft mouse models (Notta et al. 2011), underlying the fact that the HSC pool is still a heterogeneous compartment.

A very recent publication identified and characterised a prospectively separable $CD33^+$ $CD90^+$ subset of cells within the most primitive human CB cell phenotype thanks to a single cell multiparameter analysis of $CD49f^+$ HSCs. This subset showed durable repopulation potential but diverse differentiation profiles and a deeper quiescence state than $CD33^-$ HSCs (Knapp et al. 2018). More and more studies are now looking for other surface markers that may be able to isolate even purer HSCs populations and new strategies for the isolation of these cells may become soon available.

1.1.2 The haematopoietic niche

In the adult, the bone marrow is considered the primary tissue for haematopoiesis. In order to maintain the HSC pool for the lifetime of an individual, HSCs reside in a protected and highly regulated microenvironment called the BM niche. The BM niche has long been studied in mouse and a better understanding of the niche environment and molecular regulation is necessary to understand HSC maintenance and development of *ex-vivo* protocols for HSC expansion which would be extremely beneficial for HSCTs. The concept of "niche" was first proposed in 1978 by Ray Scofield and its location was later demonstrated to be in the endosteal region of the trabecular bone (Calvi et al. 2003; J. Zhang et al. 2003). However, more recent studies reported that most HSCs reside in the central marrow in proximity of sinusoids that constitute the venous circulation of the bone. These studies identified a second niche referred to as the "vascular niche".

The real location of the niche in the BM has long been debated but recent evidence shows that several niche compartments coexist to regulate HSCs in different states (Itkin et al. 2016). The BM vasculature has emerged as a key player for HSC maintenance and the presence of distinct cells surrounding BM blood vessels (perivascular microenvironment) is vital for HSC maintenance and differentiation. Therefore, the study of the niche microenvironment is important to better understand HSC regulation (Birbrair and Frenette 2016). Quiescent HSCs, characterised by low reactive oxygen species (ROS^{low}) are found to localise near the endosteal region by perivascular, endothelial, Schwann, and sympathetic neuronal cells that promote their maintenance through secretion of CXCL12 and SCF. On the other hand, activated HSCs (ROS^{high}) are found close to sinusoids in the central marrow (Itkin et al. 2016; Zhao and Li 2016). The endosteal niche is thought to provide a hypoxic environment which allows quiescence maintenance while the vascular niche regulates HSC maintenance and differentiation. Several studies investigated the specific role of different cellular components of the BM niche and identified a key role for perivascular and endothelial cells in HSC regulation within the vascular niche

(Figure 1) (Méndez-Ferrer et al. 2010; Omatsu et al. 2010; Ding et al. 2012; Pinho et al. 2013; Itkin et al. 2016). Osteoblasts were the first cells to be tested for their role in HSC regulation and, despite these cells can produce growth factors that would support haematopoietic progenitor cells like G-SCF, their role as direct regulators of HSC function has been confuted by the fact that conditional deletion of CXCL12 and SCF produced from osteoblasts doesn't affect HSCs numbers in the BM (Ding and Morrison 2013). On the contrary, perivascular cells have a central role in HSC regulation within the vascular niche. Stromal Nestin-GFP$^+$ cells physically associate with HSCs and are involved in their maintenance through SCF and CXCL12 secretion (Méndez-Ferrer et al. 2010). Nestin-GFP$^+$ also overlap with other perivascular cells like CAR cells (CXCL12-abundantt reticular cell) and LepR$^+$ cells (leptin receptor-expressing cells) which have been demonstrated to promote HSC maintenance through cytokine production (Omatsu et al. 2010; Ding et al. 2012; Pinho et al. 2013) (Figure 1). Endothelial cells also have a role in promoting HSC proliferation through expression of Notch ligands like Jag-1 and endothelial cell permeability may also be involved in the regulation of HSC quiescence and activation, however, to date, the function of single endothelial cells in the vascular niche is still poorly understood (Itkin et al. 2016). Other cells like megakaryocytes (MEGs), macrophages and non-myelinating Schwann cells were also confirmed to be involved in maintaining quiescence in HSCs and facilitate their retention within the niche (Figure 1) (Bruns et al. 2014; Zhao et al. 2014; Nakamura-Ishizu et al. 2014). Despite advances in the understanding of the niche microenvironment, none of the niche cellular components identified has appeared to be sufficient for HSC maintenance. Additional work still needs to be done in order to better understand the complexity of the niche microenvironment and how this influences haematopoiesis and maybe reveal the "secret" of HSC *ex-vivo* expansion.

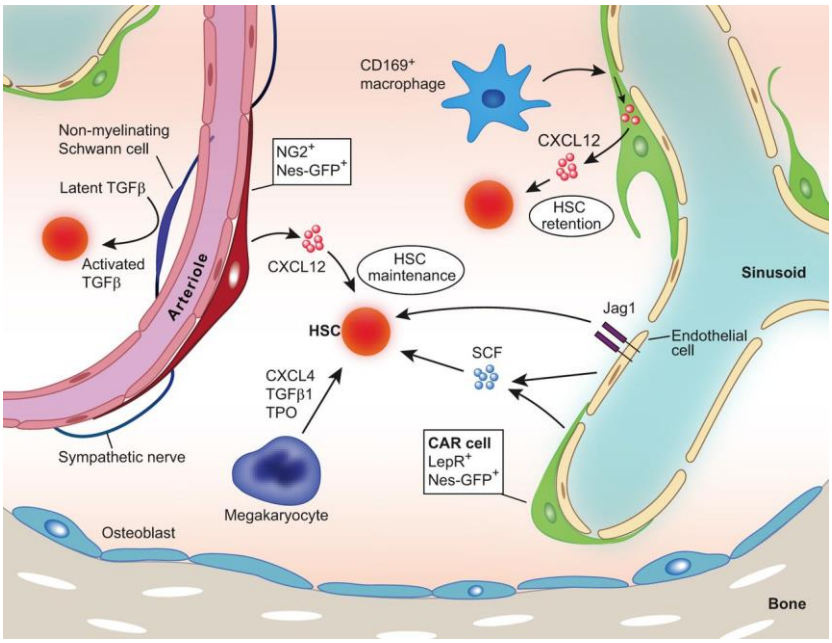


Figure 1: Schematic representation of the regulatory microenvironment of the adult BM niche and its cellular components.
Cell types identified in the BM niche and their action on HSC maintenance are shown (Gao et al. 2018).

1.2 Levels of heterogeneity in the human HSC pool

To secure their maintenance over the lifetime of an individual, HSCs are quiescent and thus reside in the G_0 phase of the cell cycle. HSCs do however divide, although very infrequently. They can commit to different fate choices. They can exit and re-enter quiescence, they can self-renew to produce identical cells or proliferate and differentiate into mature blood cell lineages (Seita and Weissman 2010). Alternatively, they can also undergo apoptosis (Figure 2).

With the extensive study of molecular and functional properties of HSCs in the mouse and later in the human, it has become clear that this compartment comprises distinct subsets of cells with differences in their capacity of self-renewal, cycling properties or differentiation potential. This introduced the concept of heterogeneity in the HSC compartment, which is now identified as a heterogeneous population composed by sub-populations with different properties. Moreover, findings demonstrating the existence of lineage-biased HSCs and lineage-restricted progenitors, challenged the classical model of haematopoiesis suggesting that lineage specification occurs much earlier than previously thought. Both cell intrinsic and extrinsic mechanisms are thought to be responsible for this functional heterogeneity (Copley and Eaves 2013; Wilkinson and Göttgens 2013). This brought the field to reconsider the classical model of haematopoiesis.

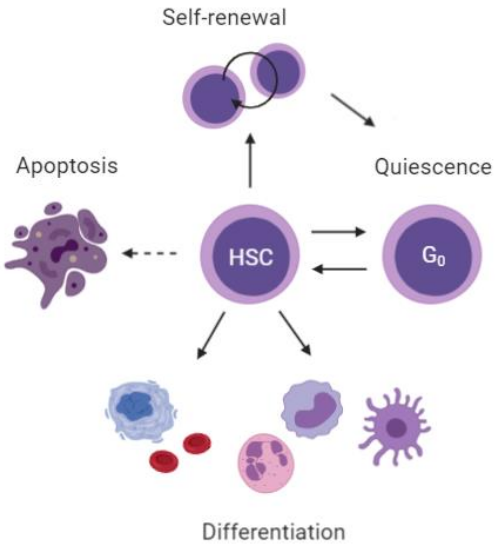


Figure 2: Summary of HSC fate choices.
HSCs can self-renew to maintain themselves, they can enter and exit quiescence, differentiate or undergo apoptosis.

1.2.1 Heterogeneity in self-renewal

In the past decade, knowing that the HSC pool is not a homogeneous stem cell compartment , different studies tried to identify possible HSC subsets based on their blood reconstitution properties and durability of engraftment following *in-vivo* transplantation, as a readout of different self-renewal properties. Single cell transplants of HSCs showed a variety in the kinetics of blood reconstitution and different self-renewal potential (Müller-Sieburg et al. 2002; Morita, Ema, and Nakauchi 2010).

Two main different sub-populations have been defined in the mouse HSC pool based on their self-renewal properties: Long-Term HSCs (LT-HSCs) and Short-Term HSCs (ST-HSCs). LT-HSCs are capable of long-term lympho-myeloid blood reconstitution for a minimum of 16 week in primary transplants (with at least 1% graft size) and they also retain the ability to engraft in secondary transplants. In contrast, ST-HSCs are defined as HSCs with more limited self-renewal and therefore only capable of short-term (more than 4 weeks but less than 16 weeks) transient engraftment after transplant (Uchida et al. 2003; Dykstra et al. 2007; Copley, Beer, and Eaves 2012). A third intermediate HSC population was identified as a subset of LT-HSCs by Benveniste et al. and defined as HSCs with intermediate reconstitution potential (IT-HSCs). The self-renewing ability of these cells is longer than ST-HSCs but shorter than LT-HSCs (Benveniste et al. 2010). Evidence in the literature reported that HSC self-renewal is affected by cell cycle regulators, transcription factors (TFs) and epigenetic regulators, however it is still not clear how this intrinsic regulation is integrated with other extrinsic factors.
On the other hand, studies on human HSCs discovered that the loss of CD90 expression was associated with loss of long-term self-renewal, demarcating a population with transient engraftment abilities (Majeti, Park, and Weissman 2007). In the attempt to identify new markers to separate HSCs from their nearest progeny, the CD49f marker was found to be expressed on ~50% of $CD90^+$ human HSCs and on ~25% of $CD90^-$ HSCs. This allowed to define in the human HSC pool the presence of a LT-HSCs sub-population identified as $Lin^- CD34^+ CD38^- CD45RA^- CD90^+ CD49f^+$ (hereby referred as $CD49f^+$ HSCs), which provides robust multilineage repopulation beyond 30 weeks while retaining serial transplantation ability, and a ST-HSCs sub-population defined as $CD34^+ CD38^- CD45RA^- CD90^- CD49f^-$ (hereby referred as $CD49f^-$ HSCs), which can generate multilineage engraftments over intermediate periods but lacks serial transplantation ability (Notta et al. 2011). The transcriptional analysis of these two cell types confirmed a high degree of transcriptional similarity (Laurenti et al. 2013) but further studies are needed to better define the mechanisms underlying these differences in self-renewal properties.

1.2.2 Heterogeneity in division and cell cycle properties

A defining feature of adult HSCs is the ability to reside outside the cell cycle in a dormant state called quiescence or G_0 state. Studies in adult mice indicate a level of heterogeneity in the division properties of HSCs, suggesting that they are not uniformly quiescent. Indeed, two different populations have been defined according to their division kinetics and they have been named “active” and “dormant” HSCs (A. Wilson et al. 2008). *In-vivo* label incorporation-based assays using DNA analogues such as BrdU (5’-bromo-2’-deoxyuridine) or histone 2B-GFP retention analysis identified a more dormant sub-population in the HSC pool, based on their degree of quiescence. This sub-population differs in the frequency of division and it accounts for the 5-10% of the HSC pool. When compared to the other 90-95% of less quiescent HSCs, the highly dormant HSC pool shows a turnover of 145 days *in-vivo* while the more active counterpart divides approximately every 30-50 days and it is thought to contribute more actively to the continuous blood formation. Nonetheless, it has been shown that despite being in a deep

quiescent state the most dormant cells display the longest repopulation capacity upon transplantation and they contribute as a reservoir in case of injury, but they are not involved in everyday blood cells production (A. Wilson et al. 2008; Foudi et al. 2009). In human, the rate of division of the most dormant human HSCs was calculated, thanks to the use of X-chromosome inactivation ratios to track HSC replications, to be around once every 40 weeks (Catlin et al. 2011). These findings support the idea that infrequent cycling of HSCs is a protective mechanism adapted to preserve genome integrity avoiding the accumulation of DNA damage due to replicative and oxidative stress.

Studies by Elisa Laurenti confirmed that there is heterogeneity in the division properties also in the human HSC pool, and they highlighted differences in stem cell quiescence states between LT-HSCs and ST-HSCs. The division frequency of human LT- and ST-HSCs was assessed by BrdU incorporation in xenograft transplantation assays and, similarly to mouse models, LT-HSCs were found to divide less frequently than ST-HSCs. Moreover, the transcriptional analysis of these populations conducted at different times after xenotransplantation identified 241 genes differentially expressed, indicating that LT- and ST-HSCs have distinct transcriptional profiles with most of the differences in genes involved in the regulation of the cell cycle (Laurenti et al. 2015). Table 1 summarises the main differences in surface markers expression, repopulation capacity self-renewal and division frequency between LT- and ST-HSCs.

Population	**Surface markers**	**Repopulation capacity**	**Serial transplantation**	**Division frequency**
LT-HSCs (CD49f$^+$ HSCs)	Lin$^-$ CD34$^+$ CD38$^-$ CD45RA$^-$ CD90$^+$ CD49f$^+$	Long term	YES	Rare
ST-HSC (CD49f- HSCs)	Lin- CD34$^+$ CD38-CD45RA- CD90$^+$ CD49f$^-$	Short term	NO	Often

Table 1: Summary of the differences between LT-HSCs and ST-HSCs.

The analysis of the cell cycle status on freshly isolated LT- and ST-HSCs showed that 90% of cells from both subsets are quiescent as defined by Ki67 staining. However, the differences found in the expression of cell cycle genes suggested a possible difference in cell cycle regulation between LT- and ST-HSCs. Molecular analysis of these two HSC subsets revealed that ST-HSCs have higher CDK6 mRNA and protein expression than LT-HSC. CDK6 is an important regulator of the entrance and progression in the G_1 phase of the cell cycle. On the other hand, LT-HSCs lack CDK6 protein, showing that ST-HSCs, despite being equally quiescent are in a CDK6-primed quiescent state that confers them a competitive advantage in shortening the duration of exit from quiescence (Laurenti et al. 2015). CDK6 expression levels were then found to regulate the duration of the quiescence exit upon mitogenic stimulation in the human HSC pool underlying a new level of heterogeneity in HSC quiescence states (Figure 3). More recently, also retinoic acid (RA) signalling was identified as important in the regulation of HSC dormancy and prolonged self-renewal together with low Myc expression in mouse (Cabezas-Wallscheid

et al. 2017). However, what other molecular mechanisms, beside RA and CDK6 levels, underlie the regulation of quiescence exit in this compartment is still unknown.

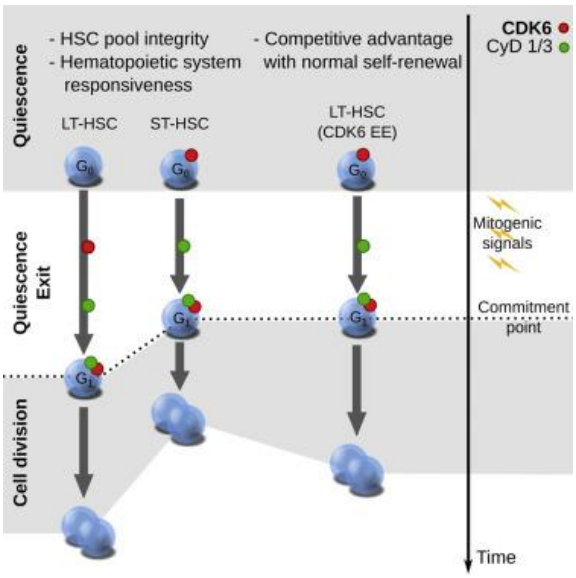


Figure 3: The haematopoietic stem cell compartment is heterogeneous in term of cell cycle properties.
Summarising scheme of the differences in the division kinetics, expression of CDK6 and duration of quiescence exit between human LT- and ST-HSCs (from Laurenti et al., 2015).

1.2.3 Heterogeneity in repopulation potential and lineage commitment

Single cell transcriptomic analyses, *in-vitro* differentiation assays and single HSC transplants have revealed another level of heterogeneity related to HSC repopulation potential and lineage commitment. Initially, limiting dilutions transplants in mouse showed that different HSC behaviours can be inherited by daughter cells and are due to intrinsic differences in self-renewal and differentiation properties (Muller-Sieburg et al. 2004, 2012). Subsets of HSCs with different repopulation potential have thus been identified and characterised as "myeloid-biased" (My-bi), "lymphoid-biased" (Ly-bi) and "balanced" (Bala) based on the ratio of lymphoid to myeloid lineage detected in donor derived blood cells after 20 weeks from transplantation (Morita, Ema, and Nakauchi 2010; Muller-Sieburg et al. 2012). Another study from Eaves et al. provided a similar categorisation using a different method that measured single HSC relative contribution to the total myeloid and lymphoid circulating cells (Dykstra et al. 2007). In this case, the classification included "lymphoid-deficient" α-HSCs primarily giving rise to myeloid cells and "balanced" β-HSCs with balanced lymphoid and myeloid lineage output and both able to repopulate secondary recipients. On the other hand, single HSCs unable to be serially transplanted were categorised as "myeloid-deficient" γ- or δ-HSCs, where γ-HSCs are lymphoid-biased but can produce myeloid cells, whereas δ-HSCs only produce lymphoid cells (Dykstra et al. 2007). According to Benveniste et al. γ- or δ-HSCs coincides with IT-HSCs (see chapter 1.2.1) since they show prolonged repopulation capacity in between of ST- and LT-HSCs (Benveniste et al. 2010). Prospective isolation of these functionally heterogeneous populations has been possible thanks to the identification of new markers including

CD150, whose expression correlates with myeloid bias and higher self-renewal potential. In fact, a higher expression of CD150 ($CD150^{high}$) correlates with α- and β-HSCs enrichment while lower levels of CD150 ($CD150^{low}$) correlate with γ- or δ-HSCs (Kent et al. 2009; Morita, Ema, and Nakauchi 2010). Despite the differences in the nomenclature of two classifications proposed above, lots of similarities are shared and HSC differentiation potential seems to be linked to their reconstitution time. In a more recent review, a new classification was proposed which divided HSCs into ST- IT- and LT- based on the length of blood reconstitution in transplanted mice and lineage balance. In this classification My-bi HSCs and α- HSCs overlap with LT-HSCs, while Ly-bi HSCs and γ- or δ-HSCs overlap with ST-HSCs. IT-HSCs instead correlate with Bala HSCs and β-HSCs (Ema, Morita, and Suda 2014). For this purpose, the consideration of long-term lineage reconstitution has been reconsidered and in this new classification LT-HSCs reconstitution time is more than 12 months after transplantation, while for IT-HSCs and ST-HSCs is more than 6 and less than 6 months respectively (Ema, Morita, and Suda 2014).

Other than myeloid and lymphoid biased HSCs, recent studies have also identified platelet-biased HSCs which precede lymphoid-biased HSCs in the haematopoietic hierarchy. Evidence of this population has also been reported in unperturbed haematopoiesis thanks to lineage tracing and transposon tagging experiments (Sanjuan-Pla et al. 2013; Yamamoto et al. 2013; Rodriguez-Fraticelli et al. 2018). Overall, in mouse, it seems that heterogeneity in differentiation potential is controlled by cell-autonomous mechanisms in individual HSCs. These differences are in fact maintained by daughter cells and probably involve epigenetic mechanism (Challen et al. 2010; Copley, Beer, and Eaves 2012). However, the molecular basis of this heterogeneity is still unclear. Although most of the identified heterogeneity seems to be intrinsically determined in HSCs, there is also a possibility that HSC differentiation and self-renewal could be influenced by external signals and modulated by the surrounding niche microenvironment. Some examples include is the responsiveness to TGF-β or IL-7 signalling which were found to be different in My-Bi and Ly-bi HSCs (Muller-Sieburg et al. 2004; Crisan and Dzierzak 2016). Moreover, niche changes during ageing and differences in the niche microenvironment (endosteal or central niches) have recently been associated with differences in HSC function. Mouse My- and Ly biased HSCs were demonstrated to be regulated by and to reside in distinct bone marrow niches respectively occupied by megakaryocytes (MK) and arterioles (Pinho et al. 2018). Moreover, age-related niche remodelling was found to be associated with My expansion (Y. H. Ho et al. 2019).

Despite the extensive evidence obtained with single cells transplantation experiments in support of HSC heterogeneity, evidence has emerged from lineage tracking studies showing that these behaviours are probably not reflective of steady-state haematopoiesis in native conditions prompting for a major role of multipotent progenitors (MPPs) in homeostatic blood production (Sun et al. 2014; Busch et al. 2015). Investigation of the MPP progenitor compartment in mouse revealed the presence of lineage-biased MPPs (MPP2, MPP3 and MPP4) which were found to work together to guarantee adequate blood production (Cabezas-Wallscheid et al. 2014). Similarly, in human, Notta et al. re-classified the MPPs compartment ($CD34^{+}$ $CD38^{-}$ $CD90^{-}$ $CD45^{-}$ $CD49f^{-}$) into three different subpopulations based on

expression of CD71 and CD110 (or BAH1) (F1 ($CD71^-$ $BAH1^-$), F2 ($CD71^+$ $BAH1^-$), and F3 ($CD71^+$ $BAH1^+$)). This uncovered heterogeneity within human MPPs and demonstrated how the hierarchy of blood cell progenitors changes across all developmental stages (fetal liver (FL), CB and BM). This showed that a higher number of cells with unilineage potential and with megakaryocytic and erythroid fates directly branches from the HSC compartment in CB and BM and it originates from the F3 $CD71^+$ $BAH1^+$ MPPs fraction (Notta et al. 2016). A summary of the HSC subsets identified to date in mouse or human, including their self-renewal potential, differentiation bias and cell cycle properties is shown in Table 2.

Species	Name	Cell-surface phenotype	Self-renewal	Cell cycle properties	Differentiation
Mouse	LT-HSC	$Lin^-Sca1^+cKit^+CD34^-CD150^+CD135^-CD48^- \pm EPCR^+ \pm Rho^{lo}$	High		
	IT-HSC	$Lin^-Sca1^+cKit^+CD34^{lo}CD135^-Rho^{lo}CD49b^{hi}$	Intermediate	Short G_0 exit	
	ST-HSC/MPP1	$Lin^-Sca1^+cKit^+CD135^-CD150^-CD48^-$	Low		
	α	NA	High	NA	Ly-deficient
	β	NA	High	NA	Balanced
	γ	NA	Intermediate	NA	My-deficient
	δ	NA	Low	NA	My-deficient
	MPP2	$Lin^-Sca1^+cKit^+CD135^-CD150^+CD48^+$	Low	Similar	Ly-deficient
	MPP3	$Lin^-Sca1^+cKit^+CD135^-CD150^-CD48^+$	Low		Balanced
	MPP4	$Lin^-Sca1^+cKit^+CD135^+CD150^-CD48^+$	Low		Ly-biased
	d-HSC	NA	Higher	Dormant	
	a-HSC	NA	Lower	Activated	
Human	LT-HSC	$Lin^-CD34^+CD38^-CD45RA^-CD49f^+CD90^+ \pm Rho^{lo}$	High	Long G_0 exit	
	ST-HSC/MPP	$Lin^-CD34^+CD38^-CD45RA^-CD49f^+CD90^+$	Low	Short G_0 exit	
	$CD34^-$ LT-HSC	$Lin^-CD34^-CD38^-CD93^{hi}$	High	Highly quiescent	

Table 2: Summary of phenotypic or functionally defined HSC subsets in mouse and human.

Mouse and human HSC subtypes identified in the literature to date are summarised here. Cell surface phenotype, self-renewal potential, cell cycle properties and differentiation potential are shown (Laurenti and Göttgens 2018).

Single cell transcriptional profiling of previously identified mouse HSC subset populations, supported the results obtained from single cell transplants revealing the expression of lineage priming modules reflective of the functional lineage biased seen *in-vivo* (Månsson et al. 2007; Pina et al. 2012; Guo et al. 2013; Moignard and Göttgens 2014; Grover et al. 2016). More recent papers, comprehensively analysed the whole transcriptional landscape of the full haematopoietic hierarchy in mouse (Nestorowa et al. 2016) or the human HSC and progenitor compartment (Perié et al. 2015; Paul et al. 2015; Velten et al. 2017; Karamitros et al. 2018). Together, these data show that lineage commitment is more of a continuous process where single cells with unilineage transcriptional profiles directly branch from a continuum of low-primed HSPCs without transitioning through multi, oligo or bipotent progenitors, thus suggesting that steps of lineage commitment may already occur in the phenotypic HSC/MPP compartments. This was demonstrated to be the case for dendritic cells (DCs) specification from human progenitor cells ($CD34^+$) which was proven to happen in parallel to myeloid and lymphoid specification thanks to specific transcriptional programs which involved expression of IRF8 in HSCs (J. Lee et al. 2017). Overall, some questions remain unanswered and when the first lineage restriction events occur or what exactly drives this heterogeneity is still not known.

1.2.4 Refining the haematopoietic hierarchy

In the past two decades, the emergence of the concept of heterogeneity in the HSC and progenitor compartments has challenged the classical view of the haematopoietic tree. Different aspects of HSC heterogeneity have been discussed above and here I will give a general overview of how the haematopoietic tree has changed from the classic model to the current one.

The classical model, also called bifurcation model, derived from the studies of Akashi and Kondo in the murine system, proposed that HSCs pass through defined steps of oligopotent, bipotent and unipotent progenitors and progressively restrict their differentiation potential in a stepwise manner to finally produce mature blood cells of different lineages. According to this model, HSCs reside at the top of the tree and are characterised by long-term repopulation potential, thus the name LT-HSCs. ST-HSCs then originate following a loss of self-renewal potential after which the first lineage restriction separates myeloid and lymphoid fates in a binary branching point that sees a separation between common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs). Megakaryocyte-erythrocyte progenitors (MEPs) and granulocyte-monocyte progenitors (GMPs) then originate from CMPs (Kondo, Weissman, and Akashi 1997; Akashi et al. 2000) while CLPs give rise to all lymphoid lineages. In this model each population is considered discrete and homogeneous (Figure 4, left).

A few years later, the identification of lymphoid-primed multipotent progenitor (LMPPs) both in mouse and human, challenged the classical model. For the first time a new multipotent progenitor population with both myeloid and lymphoid but no megakaryocyte-erythroid potential was identified (Adolfsson et al. 2005; Månsson et al. 2007; Doulatov et al. 2010). An alternative model was thus proposed where myeloid and lymphoid lineages would remain associated within LMPPs whereas progenitors of the megakaryocytic and erythroid lineage (CMPs and MEPs) would branch out closer to HSCs/MPPs, CMPs and LMPPs would then both give rise to GMPs (Figure 4, right). However, after evidence in the field suggesting that differentiation does not always occur through maturation into intermediate progenitors, another interpretation of this model was proposed. Following the discovery of cells with self-renewal capacity but megakaryocyte-restricted lineage potential in the phenotypic HSC compartment (Sanjuan-Pla et al. 2013; Yamamoto et al. 2013; Haas et al. 2015; Notta et al. 2016; Rodriguez-Fraticelli et al. 2018; Carrelha et al. 2018) this new model would include a direct shortcut from HSCs to the megakaryocytic-erythroid lineage (Figure 4, right). A major pitfall of the models discussed above is that LMPPs and other progenitor populations rely on the analysis of defined populations purified through specific cell surface markers combinations using fluorescence-activated cell sorting (FACS). This is indeed a biased method which derives information that is representative of the average of the studied population.

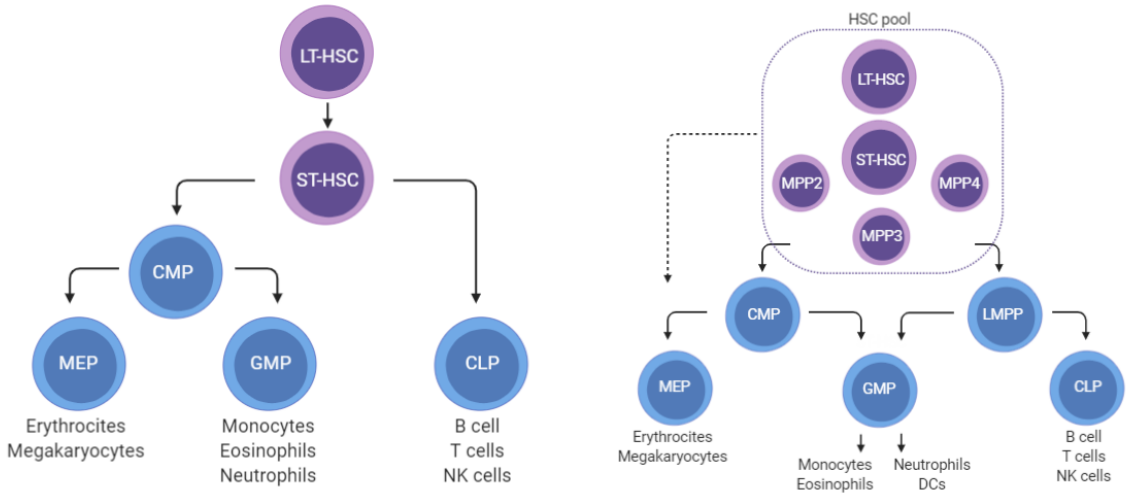


Figure 4: Graphical representations of the haematopoietic hierarchy.

The classical model, also called bifurcation model is shown on the left. The revised model which includes LMPPs is shown on the right (Based on Ema, Morita, and Suda 2014; Haas, Trumpp, and Milsom 2018; Laurenti and Göttgens 2018).

LT-HSC: long-term HSC; ST-HSC: short-term HSC; MPP: multipotent progenitor; CMP: common myeloid progenitor; MEP: megakaryocyte and erythrocyte progenitor; CLP: common lymphoid progenitor; LMPP: lymphoid-primed multipotent progenitor; GMP: granulocyte and macrophage progenitor; DC: Dendritic cell; NK: natural-killer cell.

According to recent advances in the field and thanks to extensive single cell transcriptomic analyses of HSCs, MPPs and downstream progenitors, it is now known that the frequency of unipotent progenitors is higher than previously thought and such progenitors can branch out directly from the HSC compartment (Yamamoto et al. 2013; J. Lee et al. 2017; Haas, Trumpp, and Milsom 2018). However, the unilineage potential readout doesn't necessarily mean that these cells may not have multi, oligo, or bipotential readouts in a different context (Carrelha et al. 2018). This put into discussion all the previous models and suggested that lineage choices may occur earlier than previously thought and probably already within the phenotypic HSC/MPP compartment.

In the first experiment of its kind, Velten et al. analysed human HSC lineage commitment integrating flow cytometry, transcriptomics and single cell functional data. This revealed a high degree of connection between single $CD34^{+}$ $CD38^{-}$ cells and identified a continuum of low-primed undifferentiated HSPCs (CLOUD-HSPCs) characterised by gradual differences between cells within this compartment (Velten et al. 2017). In this representation, the main route of lineage commitment is represented by direct transition from primed multilineage states towards a unilineage transcriptomic state. Distinct lineages thus emerge directly from this cloud by-passing discrete progenitors' populations. Since then, many other studies, both transcriptomic (Moignard et al. 2013; Nestorowa et al. 2016; Pellin et al. 2019; Hay et al. 2018) and functional (Perié et al. 2015; Paul et al. 2015; Velten et al. 2017; Karamitros et al. 2018; Zheng et al. 2018) supported this idea and a new haematopoietic model was suggested. In this model (Figure 5 left) HSCs do not pass through discrete intermediate states but gradually acquire intrinsic differences in lineage bias and self-renewal potential within a continuum of

transitional stages. An alternative visualisation of this model was recently proposed by Laurenti and Göttgens (Laurenti and Göttgens 2018). In this representation single HSCs move along newly identified differentiation trajectories and they transition though phenotypically defined compartments. The lineage potential of the cells in each compartment is shown and it highlights the prevalence of unilineage cells over tri and bilineage ones (Figure 5, right).

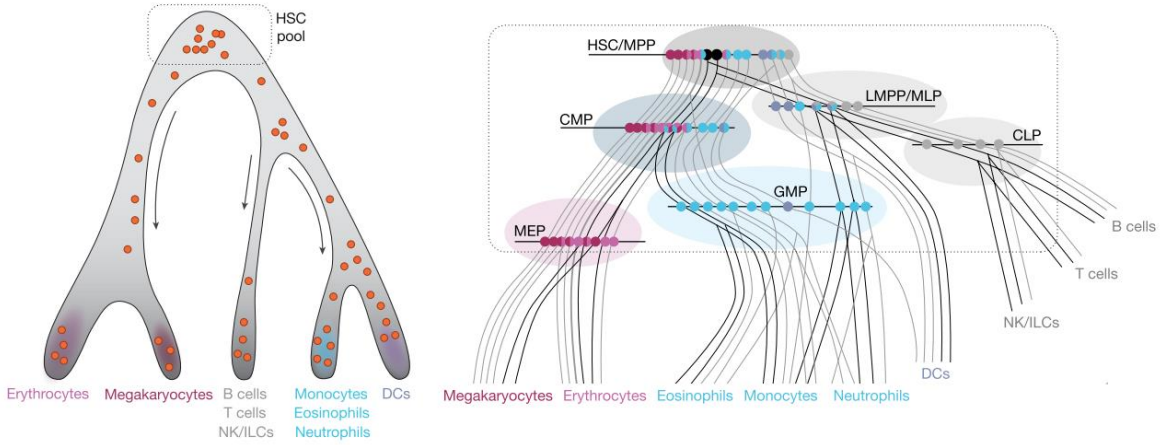


Figure 5: New hierarchical models of haematopoiesis.

The model of continuum of differentiation is shown on the left. Red dots represent single cells and their position is indicative of their differentiation trajectory. The trajectory-based visualisation suggested by Laurenti and Göttgens is shown on the right. Trajectories are shown as single lines and shaded areas represent phenotypic HSC compartments. Snapshot of lineage potential are represented by horizontal lines. The number of colors is inficative of the lineage potential (single colour= unilineage; two colours=bilineage; three colours= trilineage; black= multipotent) (Laurenti and Göttgens 2018).

1.3 Cellular and molecular properties of HSCs

The study of the cellular and molecular properties of HSCs started a few decades ago in mouse. Molecular and biochemical pathways involved in the regulation of HSC function and important for the regulation of the balance between self-renewal and differentiation were initially investigated. Gene knock-outs and enforced expression, together with transplantation experiments were predominantly used to identify TFs involved in HSC function, formation and differentiation (Rossi et al. 2012). Most of these studies were based on the classic model of haematopoiesis and thus relied on the prospective purification of defined HSC and progenitor populations. Several studies defined and categorised TFs involved in HSC maintenance and lineage specification (Orkin and Zon 2008). However, it must be considered that the mechanisms regulating cell fate choices are not driven by single TFs but are the result of a complex crosstalk between different factors. In mouse, most of the regulatory maps involved in HSC lineage commitment have been identified and are being gradually refined. Because of the heterogeneity and complexity that characterise the HSC pool it has become now necessary to study these regulatory networks at the single cell level.

Single cell transcriptomics further enhanced the level of resolution in the study of HSC and progenitors molecular properties and transcriptional states but also confirmed the heterogeneity in TF expression within these populations (Moignard et al. 2013; Moignard and Göttgens 2014; N. K. Wilson et al. 2015; Paul et al. 2015; Nestorowa et al. 2016; Hamey et al. 2017). Bioinformatical integration of single cell expression data allows to derive regulatory connections between TFs with the possibility of reconstructing network hierarchies involved in stem cell fate choices (Hamey and Göttgens 2019). An example of this comes from the study of Moignard et al. where the expression analysis of a network of TFs in more than 500 single HSPCs identified a new regulatory connection between Gata2 and Gfi1/Gfi1b during myelo-lymphoid specification where the antagonism between Gfi1 and Gfi1b is modulated by Gata2. In this case, the downregulation of Gata2 and Gfi1b by Gfi1 represent a key event in lymphoid lineage specification (Moignard et al. 2013). In another study by Cabezas et al., a genome wide molecular characterisation of highly purified HSCs and MPPs from mouse BM is obtained through quantitative proteomics, transcriptomics and methylome analysis and then validated through *in-vitro* and *in-vivo* functional assays. Here, they identify progressively changing landscape in cell type-specific methylation, gene and protein expression but also a high number of alternatively spliced transcripts isoforms which were not previously identified in these populations (Cabezas-Wallscheid et al. 2014).

Looking instead at the human counterpart, initial gene expression profiling experiments conducted on human HSCs identified some degree of conservation but also specific differences between mouse and human transcriptional profiles (Ivanova et al. 2002; Georgantas et al. 2004). Thus, due to these species-specific differences, it is important to understand how molecular programmes are associated with HSC differentiation directly into human samples, as this is crucial for the development of new clinical therapies. The molecular characterisation of human HSCs, similarly to what happened in mouse, was initially based on the prospective purification of HSCs and progenitor populations accordingly to the classical model of the haematopoietic hierarchy. MLL, RUNX1, SCL/TAL1 and LMO2 were just some of the TFs identified as involved in HSC formation and maintenance and they were also commonly found in translocations of leukaemic patients (Orkin and Zon 2008). A genome-wide expression analysis was conducted in 38 human hematopoietic cell subtypes and it identified the main TF circuits active in stem, progenitor and differentiated cells. The expression of several gene modules was found restricted to specific lineages while others were expressed over multiple lineages (Novershtern et al. 2011). The identification of a population of early-lymphoid biased progenitors that retains myeloid potential called "multi-lymphoid progenitors" (MLPs) challenged the view of a unique binary divergence between myeloid and lymphoid fates (Doulatov et al. 2010). The transcriptional dynamics of early-stage human haematopoiesis were thus investigated with a focus on the understanding of the transcriptional programs underlying lymphoid-versus-myeloid lineage choices in primary human HSCs. The gene-expression profiles of 10 prospectively isolated populations of CB HSCs and progenitors were analysed. In this study, the gene expression of MLPs clustered closely with the HSCs and not with more lymphoid-committed fractions. Differences in gene expression between multipotent and myeloid-committed

progenitors were instead clearly demarcated underlying a transcriptional dichotomy between the HSC pool and haematopoietic progenitor cells. Moreover, gene expression changes along the lymphoid lineage resulted to be more gradual with a persistent association with myeloid-associated genes (Laurenti et al. 2013). Early step of lymphoid development will be discussed more in detail in section 1.3.1.

Interestingly, as previously reported in mouse (see above), a recent study identified another level of regulation mediated by non-coding RNA and splicing variants during early stages of lineage commitment in human HSCs and progenitors. Here, the authors stress the importance of looking at cell-specific transcripts variations and not only at the gene expression level as the former ones are responsible of exerting the biological function and they do not result in noticeable gene expression changes. Novel transcript isoforms, splicing junctions and alternative splicing events were characterised at the genome-wide level and recorded providing evidence of another level of regulation during lineage specification in HSCs and progenitors (Chen et al. 2014).

More recently, the integration of flow cytometry, transcriptional and functional data at the single cell level in human BM samples demonstrated how unilineage restricted cells can originate from a continuum of low primed undifferentiated HSC and progenitor cells (CLOUD-HSCPs). In this model co-expressed gene modules associated with lineage priming were identified and used to demonstrate how distinct lineages can derive directly from lineage-primed HSCs without passing through discrete progenitor populations (Velten et al. 2017). The expression of HOXA3, PRDM16 and HOXB6 was associated with typical HSC properties such as quiescence and low RNA content, representing the least-primed state. Common HSC specific genes like HLF, MECOM and GATA3 were also found highly expressed in immature HSCs but their expression was also present in the entire 'CLOUD'. Moreover, genes modules associated with lympho-myeloid or megakaryocyte-erythrocyte priming including genes such as FLT3 or GATA2 were already expressed in HSCs demonstrating how these cells can express stem cell modules together with early lineage priming modules.

Other than TF regulation for HSC differentiation and maintenance, HSC metabolism and mitochondria function play a very important role. Cell cycle and quiescence are also strictly regulated and because of the importance of this in my research, these mechanisms will be discussed more in detail in the next sections.

1.3.1 Steps of lymphoid development

A brief overview of the molecular properties of HSCs was given in the previous paragraph. The aim of this paragraph is to dig a little bit deeper into the steps of lineage commitment with a specific focus on the steps of lymphoid lineage development. As discussed before, the haematopoiesis field has recently seen a revolution in the analysis of transcriptional networks and differentiation maps, thanks to the advent of single cells transcriptomics. In 2001 a new "myeloid-based" model was proposed by Kawamoto et al. in which lymphoid and myeloid fates would not separate early on in the haematopoietic

hierarchy but they would rather remain coupled (H Kawamoto et al. 1999; Hiroshi Kawamoto et al. 2010). A new population defined as lymphoid-primed multipotent progenitors (LMPPs) was thus identified in mouse BM and defined as Lin^{-} $Sca\text{-}1^{+}$ $c\text{-}kit^{+}$ $CD34^{+}$ $CD135^{+}$ (Adolfsson et al. 2005).

Similarly, in human, two lympho-myeloid populations were identified, both lacking erythroid and megakaryocytic potential. Human LMPPs retaining B cell, T cell, granulocytic and monocytic potential were identified as Lin^{-} $CD34^{+}$ $CD38^{-}$ $CD90^{-}$ $CD45RA^{+}$ $CD10^{-}$ (Goardon et al. 2011). On the other hand, multi-lymphoid progenitors (MLPs) defined as Lin^{-} $CD34^{+}$ $CD38^{-}$ $CD90^{-}$ $CD45RA^{+}$ $CD10^{+}$, retained lymphoid, monocytic and dendritic potential but were unable to produce granulocytes (Doulatov et al. 2010; Laurenti et al. 2013). MLPs are largely lymphoid restricted and express the early B cell marker CD10. The identification of these two populations facilitated the study of the molecular mechanisms driving myeloid versus lymphoid lineage choices in human HSCs.

In one study, MLPs transcriptomic landscapes were compared to HSC and other progenitor populations (MEPs, CMPs, GMPs, ETPs, ProB and NK cells). This showed that the transcriptional architecture of MLPs is more similar to the one of HSCs suggesting that, similarly to the mouse, lymphoid specification happens more gradually and it is not a binary decision (Laurenti et al. 2013). When looking at specific transcriptional programs, MLPs showed enrichment for gene signature specific for both myeloid and lymphoid lineages and in part of HSCs, demonstrating that HSC programmes are not totally switched off in MLPs. Among the TFs identified in MLPs specification, TFs like C/EBP-α, IKFZ1 and EBF1, important for both lymphoid and myeloid commitment were included. Moreover, other factors like BCL11A, SOX4 and TEAD1 emerged as important for the specification of the lymphoid lineage upstream of other important regulators involved in B cell commitment (Laurenti and Dick 2012; Laurenti et al. 2013). Similarities between LMPPs and MLPs have been reported and the progressive acquisition of CD10 expression has been linked with loss of myeloid potential (Ichii et al. 2010), however, MLPs population negative for CD10 have also been identified and they resulted to have similar differentiation potential although with lower myeloid differentiation capacity (Farlik et al. 2016). Moreover, other lympho-myeloid populations have been identified in the more mature Lin^{-} $CD34^{+}$ $CD38^{+}$ $CD45RA^{+}$ compartment. One is the $CD62L^{hi}$ $CD10^{-}$ population (Kohn et al. 2012) which retains lymphoid, monocytic and DC potential and the other is the granulocyte-monocyte progenitor population (GMPs, $CD19^{-}$ $CD34^{+}$ $CD38^{+}$ $CD10^{-}$ $CD7^{-}$ $CD45RA^{+}$) which still has residual lymphoid potential (Goardon et al. 2011).

Recent studies reported a better *in-vivo* and *in-vitro* characterisation of the lympho-myeloid populations in CB and demonstrated that LMPPs, MLPs and GMPs are heterogenous populations which differ functionally and transcriptionally. While erythroid and megakaryocytic fates seem to originate directly from HSCs or MPPs (Pietras et al. 2015; Grover et al. 2016) lymphoid progenitor populations appear to be part of a continuum of lympho-myeloid progenitors. MLPs and LMPPs showed similar expression of several lymphoid-specific genes like IL-7R, LCK, SYK, ADA, HLX, LST1 and ITGAL and they only had a few differentially expressed genes, supporting the idea that these populations are very similar.

GMPs, despite being more like other myeloid progenitors such as CMPs, still retained similarities with LMPPs with less than 200 differentially expressed genes (Karamitros et al. 2018). Moreover, genes signatures enrichment analysis showed that a balance between myeloid and lymphoid signature was enriched in LMPPs including myeloid-associated genes such as TRPM2, S100A8, PADI4 and ALOX15B and lymphoid associated genes like ETS1, EBF1 and CYTIP. On the contrary MLPs showed higher enrichment for lymphoid genes signature and GMPs for myeloid associated genes. Overall, heterogeneity was found in these progenitors populations suggesting the presence of intermediate populations as part of a continuum of lympho-myeloid progenitors (Karamitros et al. 2018). To date, LMPPs/MLPs are believed to be the cells where the first lineage commitment choice towards the lymphoid branch is made, and myeloid differentiation is believed to be the default commitment program that is progressively shut down for other lineages specification.

1.3.2 Role of mitochondria and metabolism in HSCs

Emerging evidence in the field has suggested that mitochondria and metabolism in HSCs have central roles in determining HSC activity and fate. Significant progress could be obtained for HSC maintenance *in-vitro* and transplantation through the study of HSC metabolism.

In the BM, HSCs reside in a hypoxic environment and they have therefore adapted to live in hypoxic conditions, so they must adjust to respond to higher oxygen levels once they are exposed to it. Moreover, it is known that hypoxia can promote HSC quiescence and this is essential for HSC function (Hermitte et al. 2006; Shima et al. 2010). However, as cells require more energy to proliferate, it seems logical that HSC metabolism would change significantly during cell activation and differentiation *in-vivo*. Several studies demonstrated how HSC metabolism differs from downstream progenitors. HSCs generate energy through anaerobic metabolism relying on glycolysis rather than oxidative phosphorylation. This is necessary to maintain a quiescent state and a dysregulation of this metabolism would affect HSC maintenance and self-renewal (Simsek et al. 2010; Suda, Takubo, and Semenza 2011). This unique metabolic balance allows to maintain low levels of ROS, which would otherwise impair HSC function and promote differentiation (Kobayashi and Suda 2012; Ahlqvist, Suomalainen, and Hämäläinen 2015). Autophagy and other regulatory molecules like FOXO , BMI1 and ATM are also some of the players involved in ROS regulation in HSCs (Ludin et al. 2014).

The anaerobic metabolism that relies on glycolisis, characteristic of quiescent HSCs, is regulated by HIF-1α (Hypoxia Inducible Factor-1 α), a TF activated in response to reduced levels of oxygen (Semenza 2011). HIF-1α is highly expressed in quiescent HSCs and its deletion in mouse HSCs results in loss of quiescence and decreased repopulation ability mainly due to the incapacity to switch from aerobic towards anaerobic metabolism (Takubo et al. 2010; Suda, Takubo, and Semenza 2011). HIF-1α expression in HSCs was also found to be regulated by MEIS1, a stem-cell associated factor expressed which is downregulated upon differentiation. MEIS1 loss has been associated with HIF downregulation, increased ROS production and loss of quiescence (Simsek et al. 2010). This evidence supports the fact

that the unique metabolism of HSCs is not only consequence of their hypoxic environment they reside in, but an important property of quiescent stem cells.

Consistently with their reliance on anaerobic metabolism, HSCs have been reported to have low mitochondrial mass (as measured by dye-dependent methods) mainly sustained by active mitochondria elimination through mitophagy (K. Ito et al. 2016; T. T. Ho et al. 2017; Snoeck 2017). This was recently confuted by one study where the use of dye-independent methods revealed that HSCs and MPPs have a high mitochondrial content which appears to be even higher than downstream progenitors and mature cells. This was proven to be due to mitochondria dye (Mito-Tracker Green) active efflux mediated by mitochondrial membrane efflux pumps, which was proposed to lead to previous inaccurate interpretations. Here, they also demonstrated that despite their high mitochondrial mass, HSCs showed reduced mitochondrial turnover and respiration (de Almeida et al. 2017). Overall, this study highlights the importance of coupling dye-independent methods to classical mitochondrial dye staining to avoid mis-interpretation due to dye efflux mediated by influx/efflux pumps present on the mitochondria.

Although the main function of mitochondria is to maintain cellular respiration, relevant literature has shown how the role of mitochondria in HSCs goes beyond the only respiratory function. Recent studies demonstrated the role of mitochondria in calcium (Ca^{2+}) homeostasis and identified Mitofusin2 (Mfn2), a mitochondrial fusion protein, as an important player for the maintenance of HSCs with lymphoid potential thanks to its effect on enhancing the buffering of intracellular calcium (Luchsinger et al. 2016). In the same study, Ca^{2+} levels were also linked with HSC functional heterogeneity. Mouse $CD150^{low}$ HSCs, which are enriched with lymphoid biased HSCs, display lower Ca^{2+} levels than $CD150^{high}$ HSCs, which are enriched with more myeloid biased HSCs (Kent et al. 2009; Beerman et al. 2010; Hock 2010). The same group also recently demonstrated that HSCs have lower Ca^{2+} than progenitor cells and their maintenance is strongly enhanced by culture in low Ca^{2+} media. Ca^{2+} efflux pumps require glycolysis and low Ca^{2+} has a suppressing effect on mitochondrial respiration, suggesting that HSC metabolic configuration is dependent on Ca^{2+} regulation (Luchsinger et al. 2019). Recent studies demonstrated the role of the Ca^{2+}-mitochondria axis in HSC division and self-renewal. A switch in the mitochondria membrane potential is in fact required for quiescent HSCs to become activated and this is linked to the increase of the intracellular Ca^{2+} levels. On the other hand, suppression of this pathway is linked to HSC self-renewal (Umemoto et al. 2018).

Overall, in the field, the importance of mitochondria in regulating HSC function and activity is becoming increasingly recognised even though HSCs mainly rely on anaerobic metabolism (Filippi and Ghaffari 2019). Accumulating evidence favours a central role for mitochondria in HSC activation and differentiation. A study from Vannini et al. showed that modulation of HSC mitochondrial metabolism affects self-renewal and fate choices. Through *in-vivo* multi-lineage blood reconstitution assays they demonstrated that HSC characterised by lower mitochondrial activity retain long-term repopulation potential and reduction of the mitochondrial activity by uncoupling of the electron transport chain can drive HSC self-renewal through autophagy stimulation. Moreover, in homeostatic conditions and during

acute stress, the mitochondrial activity of quiescent and cycling HSCs is quite similar, suggesting that metabolic differences in HSC are indicative of fate choices rather than cell cycle states (Vannini et al. 2016). The same group also recently demonstrated that the NAD^+ boosting agent nicotinamide riboside (NR) can mediate the reduction of mitochondrial potential and activity in HSCs through increased mitochondrial clearance mediated by mitochondrial stress response and autophagy. NR would thus stimulate haematopoiesis while increasing asymmetric divisions in LT-HSCs (Vannini et al. 2019). This, once again, highlights the link between HSCs fate choices and mitochondrial regulation.

Mitochondria were also found to be asymmetrically segregated during cell division in adult stem cells and they can also be remodelled through fusion or fission (Katajisto et al. 2015; H. Zhang, Menzies, and Auwerx 2018). However, these mechanisms are still unclear in haematopoiesis and more studies are needed to better understand mitochondria's role in HSC biology. Modulation of mitochondria activity and metabolism in HSCs has important implications for their *in-vitro* maintenance and for stem cell therapies.

1.4 Cell cycle and quiescence regulation in HSCs

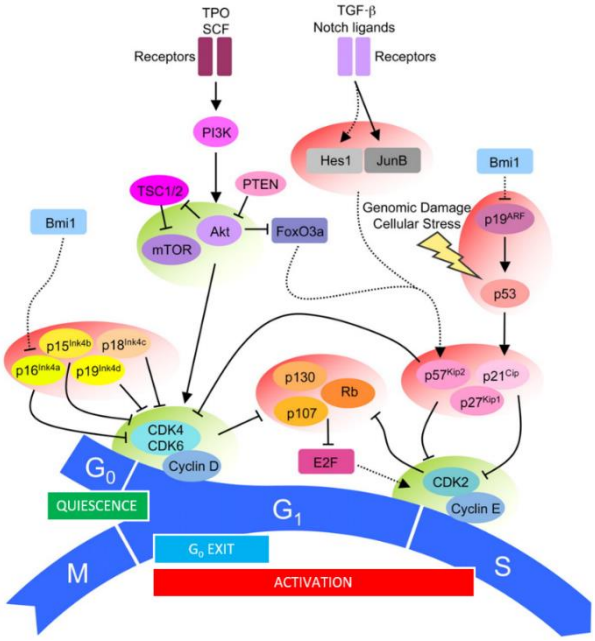
As previously discussed, an important property of HSCs is to reside in a quiescent state in homeostatic conditions. However, quiescence is a reversible state in which the cell, in response to external stimuli can decide to progress into the cell cycle. A disruption of quiescence in HSCs is highly detrimental for the organism and it leads to defects in self-renewal resulting in HSC exhaustion, as seen in knock-out models for many genes (Orford and Scadden 2008; Rossi et al. 2012). Thus, the regulation of the cell cycle in HSCs plays a critical role in the maintenance of a balance between self-renewal and differentiation. Most of the studies related to cell cycle regulation in HSCs have been conducted in mice and a detailed description of these mechanisms is still lacking in the human counterpart.

The mammalian cell cycle is classically divided into two main phases, the synthesis phase (S) and the mitotic segregation phase (M) separated by two gap phases called G_1 and G_2, preceding S and M phase respectively (Figure 7, right). A G_0 phase, or quiescent phase, has also been described as a non-cycling state outside the cell cycle. This is a reversible state and cells can re-enter the cell cycle upon sensing appropriate stimuli (Cheung and Rando 2013; Hao, Chen, and Cheng 2016). Cell cycle progression is highly regulated and the transition from G_1 to S is the most regulated step. A crucial aspect of cell cycle regulation is the existence of DNA damage and replication checkpoints at the transition from G_1 to S and from G_2 to M and the existence of a restriction point called "R point", after which a cell becomes committed to enter the cell cycle and progresses through it independently from external signals. Upon sensing errors in DNA replication or chromosome segregation, the activation of these checkpoints will arrest the cell cycle to repair possible defects. The R point, which coincides with the phosphorylation of the Retinoblastoma protein (RB), divides the G_1 phase into a mitogen-dependent early G_1 phase and a late G_1 mitogen-independent phase (Figure 6). Cyclins and cyclin-dependent kinases (CDKs) are the main regulators of the cell cycle and the kinase activity of CDK/cyclin complexes is controlled by CDK

inhibitors (CKIs), which belong to either the INK4 or CIP/KIP families (Pietras, Warr, and Passegué 2011; Hao, Chen, and Cheng 2016). CDKs are responsible for the phosphorylation of transcriptional inhibitors and they promote the activation of transcriptional waves leading to the production of important regulators of downstream cell cycle events. D-type cyclins together with CDK4 and CDK6 regulate the transition from early to late G_1 inactivating RB through its phosphorylation and thus promoting the expression of cyclin E which, together with CDK2 will hyper-phosphorylate RB confirming the entrance in late G_1 phase of the cell cycle (Figure 6) (Malumbres and Barbacid 2009; Lim and Kaldis 2013). Interestingly, while cyclin D-CDK4/CDK6 were thought to be essential for cell cycle progression, cyclin D-CDK4/CDK6 independent progression into the cell cycle has also been reported in many cellular compartments. From these studies also emerged how, instead, HSC cell cycle progression is dependent on the presence of cyclin D-CDK6/CDK4 as mice lacking these regulators presented perturbations of the hematopoietic system (Malumbres et al. 2004; Kozar and Sicinski 2005; Tigan et al. 2016).

Figure 6: Regulation of HSC cell cycle entry.

Entrance in the cell cycle is finely regulated by the interplay of cell-intrinsic and extrinsic mechanisms in HSCs. A schematic representation



of this regulation is shown here. Solid arrows indicate direct activation or inhibition events, dashed arrows indicate transcriptional regulation. Green shaded circles identify functionally related groups of cell cycle activators while cell cycle inhibitors are grouped into red shaded circles. Quiescence, exit from quiescence (G_0 exit) and activation phases as defined in this work are shown (adapted from Pietras, Warr, and Passegué 2011).

The cell cycle activity of HSCs is carefully modulated by a complex interplay of cell-intrinsic mechanisms and cell-extrinsic factors produced by the microenvironment. HSC entrance in the cell cycle corresponds to the transition from G_0 to G_1 phase, that corresponds to G_0 exit (Figure 6) and is regulated by the competition of activators and inhibitors of cyclin-CDKs complexes. Extrinsic signals important for HSC maintenance coming from the surrounding niche in the BM, like thrombopoietin (TPO), stem cell factor (SCF) or transforming growth factor β (TGF-β), contribute to the regulation of quiescence in

HSCs acting on the regulation of the cyclin D-CDK4/CDK6 complex. This complex controls the entrance in the late G_1 phase as it can phosphorylate RB thus passing the restriction point (Figure 6). RB and other members of its family (p107 and p130) of transcriptional repressors restrict cell cycle entry by repressing E2F driven transcription of positive cell cycle regulators like E-type cyclins, whose role is crucial in late G_1 and S phase. The cyclin E-CDK2 complex can then hyper-phosphorylate RB with the effect of enhancing E2F mediated transcription that will lead the progression from late G_1 to S phase (Figure 6). A firm role for RB family members in the regulation of HSC cell cycle was established after demonstrating that conditional deletion of the entire RB family in adult mice leads to an increased HSC proliferation (Viatour et al. 2008).

Moreover, the role of CKIs of the INK4 and CIP/KIP families is also important as they negatively regulate the activity of CDKs complexes. In particular, the members of the INK4 family ($p15^{INK4B}$, $p16^{INK4A}$, $p18^{INK4C}$ and $p19^{INK4D}$) function as antagonists of the cyclin D-CDK4/6 complex (Figure 6), thereby blocking RB phosphorylation and entry into S phase. Their regulation has been demonstrated to be important in order to maintain a proper balance between quiescence and proliferation in the HSC pool (Yuan et al. 2004; Oguro et al. 2006). On the other hand, the CIP/KIP ($p21^{CIP}$, $p27^{KIP1}$, $p57^{KIP2}$) family, which is highly expressed in quiescent cells, also restrain the entry into S phase but acting on the inhibition of the cyclin-E-CDK2 complex (Figure 6) *(Matsumoto et al. 2011; P. Zou et al. 2011)*. The expression of CIP/KIP family members is in turn regulated by other transcription factors including HES1, JUNB and FOXO3, which are activated by extrinsic growth-repressive signals like TGF-β. Furthermore, HSC cell cycle activity is also regulated by p53, which becomes active in response to cellular damage and thus inhibits the progression of the cell cycle favouring quiescence maintenance through a p21 independent mechanism (Liu et al. 2009a) (Figure 6). The role of all these factors in HSC cell cycle regulation has been widely validated in knock out models *in-vivo* (Pietras, Warr, and Passegué 2011). In conclusion, the regulation of the cell cycle in HSCs is quite complex and it involves the interplay of intrinsic and extrinsic factors. A better understanding of this fine regulation would give further insight on how HSCs decide to exit quiescence and progress into the cell cycle and whether this regulation is related with cell fate decisions.

1.4.1 Quiescence and activation in HSCs

While the concept of quiescence has been well described in the stem cell context, the concept of activation still lacks a clear definition as it is generally defined both as the entrance in the cell cycle or increased frequency of cycling of HSCs. Because I will base most of my work on the comparison between these two states, a better definition of quiescence and activation is needed for the purpose of this study.

Quiescence has been extensively studied over the years and it is recognised as a protective phase for long-lived adult stem cells. The balance between quiescence and proliferation in stem cell populations is tightly regulated in order to prevent exhaustion of the stem cell pool and excessive cell growth.

Quiescence was initially defined as a deeply inactive and dormant state of the cell, located outside of the cell cycle in a phase called G_0. However, an increasing number of studies has demonstrated that quiescence is a poised state, which allows a rapid cell activation and differentiation upon sensing appropriate stimuli (Nakamura-Ishizu, Takizawa, and Suda 2014; van Velthoven et al. 2019). As it is often cause of confusion, it is important to remember that quiescence and senescence are not the same thing. Quiescent cells can re-enter the cell cycle whereas senescent cells are quiescent but terminally differentiated and incapable of entering any other cell cycle phase (Terzi, Izmirli, and Gogebakan 2016; Cho et al. 2019).

Common characteristics of quiescent stem cells include low RNA content and thus reduce protein synthesis, reduced metabolism and lack of proliferation markers (Figure 7, left) (Passegué et al. 2005; Simsek et al. 2010; Cheung and Rando 2013). Quiescent HSCs rely on anaerobic glycolysis as major source of energy and they have lower mitochondrial membrane potential and thus lower mitochondrial activity (see chapter 1.3.2 for HSC metabolism). The use of glycolysis has been shown to have protective advantages for resting HSCs as it minimizes the production of DNA-damaging reactive ROS, by-products of mitochondrial oxidative phosphorylation. On the contrary, when HSCs become activated they rapidly switch from anaerobic glycolysis to mitochondrial oxidative phosphorylation to facilitate the robust energy demand associated with cell cycle progression and differentiation (Folmes et al. 2012; Yu et al. 2013; Warr and Passegué 2013).

Recent studies conducted in muscle stem cells (MuSCs) showed the existence of different depths of quiescence. These cells can cycle between two different quiescence states: the G_0 phase and a more primed G_{Alert} phase characterised by increased transcriptional activity and metabolic activity than G_0. When in the G_{Alert} phase, cells can exit quiescence and respond to stress more easily and thus contribute to tissue repair in a more efficient way (Rodgers et al. 2014). Transition from G_0 to G_{Alert} is controlled by mTORC1 activity and activation of this pathways in HSCs resulted in increased mitochondrial activity and thus transition to a G_{Alert} state (Rodgers et al. 2014). The mechanisms controlling quiescence depth are still poorly known. Two recent studies demonstrated that quiescence depth can be regulated by the activation threshold of the RB-E2F switch. First, as described above in 1.2.2, both LT-HSCs and ST-HSCs are in a quiescent state but ST-HSCs can more rapidly enter the cell cycle thanks to their higher levels of CDK6 (Laurenti et al. 2015), which controls RB phosphorylation. However, as the mTORC activity is similar between these populations, the ST-HSC primed state does not correspond to what is defined as G_{Alert} state (van Velthoven et al. 2019). During quiescence the hyperphosphorylation of RB inhibits E2F transcriptional activators which are then activated though CDK4/6 phosphorylation upon sensing mitogenic stimuli. Deeply quiescent cells then need a stronger stimulation to re-enter the

cell cycle than other less quiescent cells (Laurenti et al. 2015). A graphical representation of the cell cycle and the more recently identified quiescence cycle is shown in Figure 7.

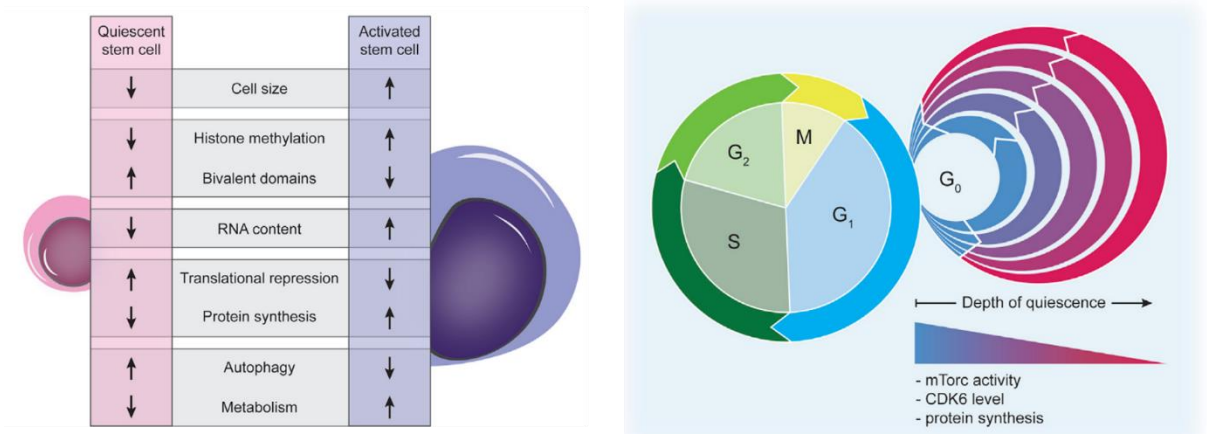


Figure 7: Characteristics of quiescent and activated stem cells and representation of the cell cycle.
A schematic comparison of the characteristics of quiescent and activated stem cells is shown on the left. Arrows are indicative of upregulation or downregulation of such processes or characteristics. A schematic representation the cell cycle and quiescence cycle is shown on the right. Different levels of quiescence are shown in correlation to mTORC activity, CDK6 levels and protein synthesis (from van Velthoven et al. 2019).

Recent studies provided better insights into the molecular regulation of quiescence suggesting that this state is maintained thanks to the activity of many different intrinsic mechanisms but also through extrinsic signals (Cho et al. 2019). Epigenetic regulation seems to play an important role in keeping quiescent cells "ready" to enter the cell cycle in other stem cells. Several genes result to be in a transcriptionally poised state which is supported by a permissive chromatin environment that will facilitate gene transcription during activation (Srivastava, Mishra, and Dhawan 2010; Kheir and Lund 2010; Cho et al. 2019). Autophagy has also been identified as important for quiescence maintenance in many adult stem cells including HSCs due to its active role in ROS reduction and removal of mitochondria (Mortensen et al. 2011; García-Prat et al. 2016; T. T. Ho et al. 2017). Genetic approaches and high-throughput analyses of several adult quiescent stem cells provided information about molecular signatures common to quiescent stem cells, including HSCs. These analyses show a down regulation of genes involved in DNA replication and cell cycle progression like Cyclins A2, B1 and E2, important regulators of the DNA checkpoints during the cell cycle. Moreover, other genes were identified as down-regulated including genes correlated with the proliferation status and mitochondrial function, whose lower expression reflects the low metabolic activity of quiescent cells (Forsberg et al. 2010; Cheung and Rando 2013; van Velthoven et al. 2017).

Due to the long life of quiescent HSCs, it is important for these cells to protect themselves from environmental stress, like ROS. The FOXO family of transcription factor has been found to play a role in HSC protection against ROS. Indeed, HSCs depleted of FOXO family members (FOXO1, FOXO3 and FOXO4) exhibit a marked increase in ROS and the propensity to exit quiescence (Tothova et al. 2007; Cheung and Rando 2013).

1.4.2 Definition of HSC activation in the context of this study

Despite recent advances in the understanding of mouse HSC quiescence, no studies have been conducted in human HSCs. Moreover, to date, there is only very limited information regarding G_0 to G_1 transition. We have defined "exit of quiescence" as the process of abandoning the quiescent status and entering the mitogen-dependent early G_1 phase right before the restriction point, which involves the phosphorylation of RB by the cyclin D-CDK4/6 complex (Figure 6). We have chosen this definition as no G_0/G_1 check points or markers have been identified to date to distinguish these two phases. In contrast to quiescence, it is still debated what the activation process consists of. Is a cell activated when it is in any of the phases of the cell cycle (G_1, S, G_2 or M) or just right after exiting quiescence? Is the activation state defined by the activation of specific signalling pathways or TFs? What are the cellular and molecular properties of an activated stem cell or HSC? None of the answers to these questions is reported in the literature. For the purpose of this study, I defined "activation" as the process of exiting quiescence and entering the cell cycle (in any of the cell cycle phases) including everything that could correlate with this action, like the activation of pathways promoting cell growth or an increase in the metabolic activity. When studying quiescence *in-vitro* it is important to remember that the removal of stem cells from their niche and their purification, which for HSCs usually requires flow cytometry sorting, already results in significant molecular and transcriptional changes within the cells. Although cell cycle analysis shows that HSCs are still quiescent, the process of activation most likely has already been initiated and the properties of such cells may not reflect their *in-vivo* state. This poses indeed challenges to the study of quiescence and quiescence exit *in-vitro* and needs to be put into consideration when interpreting results.

In conclusion, it is important to develop new strategies in order to define the molecular pathways involved in the maintenance of a quiescent status and transition to the G_1 phase in human HSCs.

1.4.3 Cell cycle and cell fate decisions

The high degree of heterogeneity found in both cell cycle and functional properties within the HSC pool suggests a possible link between the kinetics of cell cycle progression and cell fate choices. However, to date, it is still unclear if there are molecular regulators responsible of coordinating cell cycle progression and cell fate decisions in the HSC pool, including the decision to self-renew or to undergo differentiation towards a specific lineage.

Strong evidence of such coordination exists in other systems, with studies on human and mouse Embryonic Stem Cells (mESCs) and Neural Stem Cells (NSCs). These works highlighted the importance of the G_1 phase of the cell cycle in the initiation of cell fate decisions. Indeed, it has been demonstrated that, while S, G_2 and M phases lengths are comparable between different cell types, the entry and progression through G_1 varies from cell to cell. In fact, G_1 is significantly truncated in pluripotent stem cells when compared to committed or differentiated cells. mESC, and in general pluripotent stem cells, divide rapidly and spend most of their cell cycle in the S phase where they

replicate their DNA; only little time is spent in the gap phases. In these cells E- and A- type cyclins are highly expressed throughout the cell cycle and inhibition of CDKs through CDKIs is almost absent. This results in constitutive activation of CDKs which facilitates inactivation of the RB pathways and thus a rapid transition through the G_1 phase (K. A. Becker et al. 2006; Neganova et al. 2009; Boward, Wu, and Dalton 2016). Expression of CDKIs or knockdown of CDKs in human ESCs has been found to promote differentiation, suggesting that the G_1 phase may be important for fate specification (Neganova et al. 2009; Ruiz et al. 2011; Boward, Wu, and Dalton 2016).

Several studied thus identified the G_1 transition time to be important in the "decision making process", the process where the cells decides commit to a certain fate (Blomen and Boonstra 2007). For pluripotent stem cell the G_1 phase is a "window of opportunity" where they can decide to self-renew or respond more efficiently to differentiation cues. Experiments in support of this show that ESCs and NSCs are characterised by a short G_1 phase that allows them to maintain their self-renewal potential, while the lengthening of their transition time through G_1 has been associated with initiation of differentiation (Lange and Calegari 2010; Calder et al. 2013; Coronado et al. 2013). Several reports described how a short G_1 phase is an important feature of pluripotency and what is the process by which pluripotent stem cells initiate their differentiation programs in this phase (Dalton et al. 2015), but only few recent studies provided further insights in the understanding of the molecular mechanisms driving cell fate decisions during G_1 (Pauklin and Vallier 2013; Singh et al. 2013). These studies took advantage of the Fluorescence Ubiquitin Cell Cycle Indicator system FUCCI report system, which allows to sort cells in different cell cycle phases but also to monitor the length of each phase of the cell cycle without altering cellular properties (Sakaue-Sawano et al. 2008). These data revealed how important players of the cell cycle machinery contribute to a differential compartmentalisation of differentiation by driving key transcriptional factors. In one of these studies it was found that, in human ESCs, the complex cyclin D-CDK4/6 has a central role in limiting the nuclear localisation of key signal factors like SMAD2/3, whose control results in a different regulation of genes involved in endodermal differentiation. When the level of cyclins D is low, in early G_1, SMAD2/3 promotes endodermal differentiation whereas, when SMAD2/3 nuclear import is blocked due to the higher expression of cyclin D in late G_1, hESCs are only receptive for neuroectodermal initiation. This allowed to confirm that hESCs differentiation occurs during the G_1 phase and, in these cells, early G_1 and late G_1 have different differentiation capacities, implying that cell fate choices primarily occurs in early G_1 phase when cyclins D are dynamically expressed (Pauklin and Vallier 2013). Taken together, these evidences uncover the first molecular mechanisms involved in cell cycle regulated cell fate decisions in hESCs, but they also suggest that G_1 phase could be the main window for cell fate choices also in other stem cells. Moreover, this is consistent with accumulating evidence showing that CDKs display other non-canonical functions. Specifically, CDK6 has emerged as directly involved in transcriptional regulation in cancer cells and in HSCs, a role that has been demonstrated to be independent of CDK6's kinase activity (Kollmann et al. 2013). Among its many roles, CDK6 is also involved in the process of differentiation through the modulation of gene

expression changes and epigenetic status that contribute to the determination of cell fate. Evidence shows that decreased levels of CDK6 are required for erythroid differentiation (Choe et al. 2010) but also indicate that CDK6 can prevent myeloid differentiation though the binding of RUNX1 (Fujimoto et al. 2007). More CDK6 cell-cycle independent functions are being investigated and new inhibitors against the non-kinase activity of CDK6 are being considered for development (Tigan et al. 2016).

Recently, new insight emerged into mechanisms linking cell fate decisions and cell cycle control which were not only linked to the G_1 phase but also to the transition from S to G_2 phase. Gonzales et al. identified a role for S to G_2 phase regulators into control of pluripotency and differentiation of hESCs. Several epigenetic mechanisms including chromatin remodelling and histone acetylation resulted as essential for the transition from pluripotency to differentiation during S and G_2 phase. Moreover, ATR/ATM factors, which are involved in DNA damage checkpoints, were shown to enhance the TGF-β/ACTIVIN/NODAL pathway activity through p53 during the same phases (Gonzales et al. 2015; Vallier 2015).

Although hESCs and adult human HSC cell cycle dynamics are highly divergent, the delineation of a relationship between the cell cycle machinery and cell fate decision in human HSCs could provide better insights in the molecular regulation of HSC differentiation.

Recent studies in human HSCs contributed to understand how HSC function is controlled by the kinetics of cell cycle progression and in particular whether exit from quiescence rather than transition time through G_1 is the most important period for the decision of HSC fates (Laurenti et al. 2015; Mende et al. 2015). The overexpression of CDK6, a major regulator of quiescence exit, in LT-HSCs, conferred a competitive advantage for the expansion of LT-HSCs *in-vivo* without altering differentiation, contrasting with previous work reporting that increased cycling is associated with impaired LT-HSC maintenance. On the other hand, overexpressing both cyclin D1 and CDK4 in human HSCs results in a shortened transition time through G_1 associated with competitive advantage and increased myeloid differentiation (Mende et al. 2015). Collectively, these evidences suggest the idea that not only early G_1 phase but also transition from G_0 to early G_1 is a sensitive period during which cell fate decisions are taken and genes involved in the regulation of this transition phases might have an impact in cell fate decisions by altering the time spent in this sensitive period. However, in the HSC field it is still not known if cell division precedes differentiation and how these relate to each other's.

More recent *in-vivo* tracking studies in $Ki67^{RFP}$ knock-in mice followed HSC differentiation, divisional history and cell cycle control at the single cell level to identify the correlation between HSC differentiation and cell cycle control. They demonstrated that, *in-vivo*, HSCs can differentiate into restricted myeloid and megakaryocyte- erythroid progenitors in the G_0/G_1 phases and before entering the S phase of the cell cycle (Grinenko et al. 2018). They suggest that HSC division and differentiation are independent processes and fate decisions are made by HSCs before committing to divide and thus before the S phase of the cell cycle. If similar behaviours are also seen in human HSCs is not known, but they open new directions in the study of similar behaviours in human HSCs. The main goal is now

to understand if these cell cycle phases are just sensitive phases where the cell is more exposed to specific cell fates signals or if important regulators of cell cycle progression and exit from quiescence are part of a more complex picture that sees them involved in other roles independent from the cell cycle. In order to study this in HSCs, the development of methods able to identify variations in the early G_1 phase and G_0 exit and correlate these with lineage bias will be required.

2 AIMS

In the past decade, advances in single cell transcriptomic have complemented an array of functional experiments demonstrating a high degree of heterogeneity within the HSC compartment of both mouse and human. However, how this functional diversity is regulated at the cellular and molecular level and when HSC multipotency is lost is still not known.

Moreover, despite decades of studies on how HSCs maintain their quiescent status in the mouse model and more recently in human, still very little is known about how the molecular networks regulating quiescence change during the progression from quiescence to division, here termed activation. Activation has previously been associated with loss of self-renewal and initiation of differentiation programmes, but to date no studies have formally examined if cell fate decisions occur during this activation process.

In this thesis I present the work I have done in human HSCs towards two different aims:

A) Dissect human HSC heterogeneity in the purest HSC compartment ($CD49^+$ HSCs/LT-HSC) reported to date and understand when the first steps of lineage restriction occur.

B) Build a comprehensive picture of the molecular and cellular changes occurring during quiescence exit and cell cycle progression and investigate when and if human HSC self-renewal is lost upon activation.

The work and the results relative to each aim will be presented in two different chapters.

3 METHODS

NOTE: Sections of this chapter have been adapted from Belluschi et al., Nat. Communications, 2018.

3.1 Human cord blood samples

De-identified umbilical cord blood (CB) samples were obtained with informed consent from healthy donors through the Cambridge Blood and Stem Cell Biobank (CBSB) in accordance with regulated procedures approved by the relevant Research and Ethics Committees (07/MRE05/44 research study). CB units received on the same day were pooled independently of sex and processed as a single sample. CB samples used for single cell RNA-sequencing (scRNA-seq) experiments were pooled by sex (males only or females only).

3.2 Human CB CD34$^+$ cells selection

Upon receiving CB, mononuclear cells (MNCs) were isolated by Lymphoprep (Stem Cell Technologies) or Pancoll (PAN-biotech) by density gradient centrifugation of pre-diluted CB (1:1 ratio with PBS). The collected MNCs fractions were then depleted of red blood cells (RBCs) after 15 minutes incubation at 4˚C with Red Blood Cells Lysis Buffer (BioLegend). In order to proceed with the CD34$^+$ selection, a minimum number of 10^8 cells were required. If this condition was not satisfied MNCs were stored at -150˚C.

CB CD34$^+$ cells were positively selected using the Micro Beads CD34$^+$ selection kit (Myltenyi Biotec) and the AutoMACS cell separation technology (Myltenyi Biotec) according to the following adapted protocol (Table 3). CB CD34$^+$ cells were then stored at -150˚C until use.

Reagents	**Quantity**
CD34 Micro Beads	30 µl/10^8 cells
FcR Blocking Reagent	30 µl/10^8 cells
PBS + 3% Fetal Bovine Serum (FBS)	90 µl/10^8 cells

Table 3: Adapted protocol for CD34$^+$ cells selection.

3.3 Fluorescence-activated cell sorting (FACS) and flow cytometry analysis

3.3.1 Sample preparation for sorting

To isolate different cell populations from CD34$^+$ CB cells, frozen CB CD34$^+$ samples were thawed by drop-wise addition of pre-warmed Iscove's Modified Dulbecco's Medium (IMDM, Life Technologies) + 0.1 mg/ml DNase (Sigma) + 50% Fetal Bovine Serum (FBS, Life Technologies) and then re-suspended after counting at a concentration of $2x10^6$ cells/ml in PBS + 3% FBS antibody mix. Cells were then incubated 20 minutes in the dark at room temperature (RT) and washed with 10 volumes of

PBS + 3% FBS. Depending on the experiment $CD34^+$ CB cells were stained with antibodies against different panels of surface markers (Table 4) and sorted on BD FACS Aria III or on BD FACS Aria Fusion sorters available at the NIHR Cambridge BRC Cell Phenotyping Hub facility. For single cell experiments, cells were sorted using single cell purity and index sorting to allow retrospective correlation between the transcriptomic or functional information obtained from each single cell by *in-vitro* functional assays or scRNA-seq and its cell surface markers expression at the time of the sort. Purity mode was used for all bulk sorts. All sorts showed >95% purity.

The following markers were used to sort previously defined populations: HSCs/MPPs ($CD19^-$ $CD34^+$ $CD38^-$ $CD45RA^-$), $CD49f^+$ HSCs/LT-HSCs ($CD19^-$ $CD34^+$ $CD38^-$ $CD45RA^-$ $CD49f^+$ $CD90^+$), $CD49f^-$ HSCs/ST-HSCs ($CD19^-$ $CD34^+$ $CD38^-$ $CD45RA^-$ $CD49f^-$ $CD90^-$), LMPPs ($CD19^-$ $CD34^+$ $CD38^-$ $CD45RA^+$ $CD10^+$), MLP ($CD19^-$ $CD34^+$ $CD38^-$$CD45RA^+$ $CD10^-$), GMP ($CD19^-$ $CD34^+$ $CD38^+$ $CD10^-$ $CD7^-$ $CD45RA^+$) and CMPs/MEPs ($CD19^-$ $CD34^+$ $CD38^+$ $CD10^-$ $CD7^-$ $CD45RA^-$), after gating out dead cells based on Zombie Aqua (Biolegend).

Several combinations of antibodies were used depending on the type of experiment

Antibody panel combination	**Antibody (clone)**	**Dilution**	**Fluorochrome**
A, B, C	CD19 (HIB19)	1:300	Alexa 700
A, B, C	CD34 (581)	1:100	APC-Cy7
A, B, C	CD38 (HIT2)	1:100	PE-Cy7
A, B,	CD45RA (HI100)*	1:100	FITC
C	CD45RA (HI100)	1:100	PE
A, B,C	CD49f (GoH3)*	1:100	PE-Cy5
A, C	CD90 (5E10)*	1:100	APC
B	CD90 (5E10)	1:50	PE
A, B	CD10 (HI10a)*	1:100	BV421
A, B	CD7 (M-T701)*	1:100	BV421
A	CLEC9A (8F9)	1:75	PE
A, B	CD117 (104D2)	1:100	BV650
A, B,C	Zombie	1:2000	Aqua

Table 4: Antibodies and antibody panels used for isolation of cell populations derived from $CD34^+$ cells.

All the antibodies used for cell sorting are listed in this table. Several combinations of antibodies were used depending on the type of experiment. For each cell surface marker directed antibody, clone, dilution and fluorochrome are specified. All the antibodies listed were purchased from BioLegend except those indicated by * that were purchased from BD bioscience. Before use all the antibodies were titrated using appropriate control samples from human CB.

3.3.2 FACS gating strategies

The following gating strategies were used to sort previously known populations including: $CD34^+$ $CD38^+$ progenitor cells, $CD34^+$ $CD38^-$ cells, HSC pool cells, phenotypic LT-HSCs ($CD49f^+$ HSCs) and ST-HSCs ($CD49f^-$ HSCs) GMPs, MLPs, LMPPs

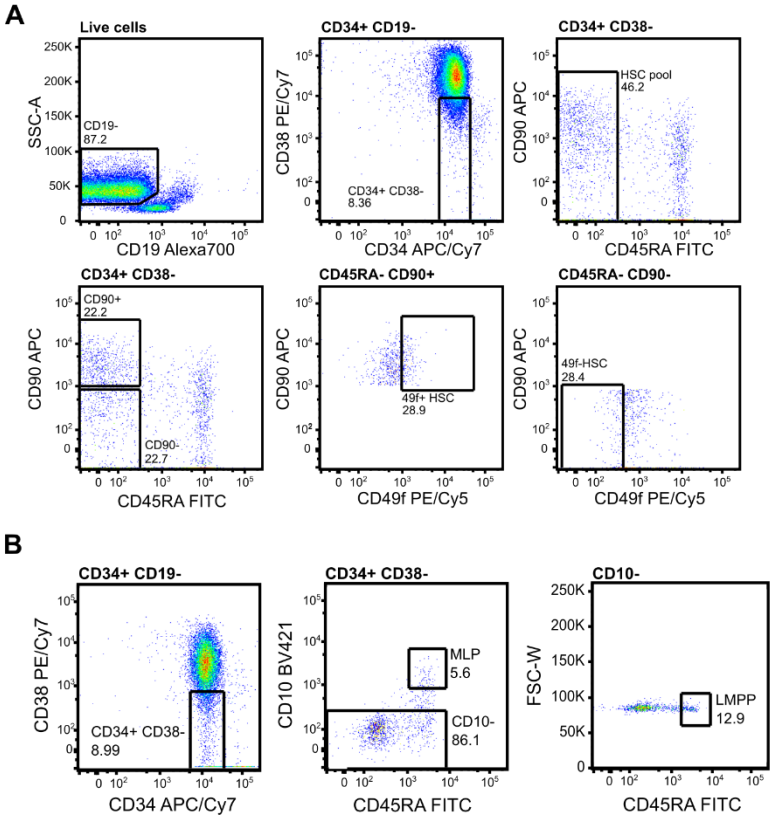


Figure 8: Gating strategy used to isolate HSC pool cells, phenotypic LT-HSCs and ST-HSCs, GMPs, MLPs and LMPPs.

A, B) FACS plots are representative examples of sorting strategies used to isolate the indicated populations. Gates and percentage of cell populations in each gate are shown. A) Representative example of gating strategy used to isolate HSC/MPPs, $CD49f^+$ HSCs and $CD49f^-$ HSCs from $CD34^+$ CB cells. $CD49f^+$ HSCs are defined as the highest 30% $CD90^+$ $CD49f^+$ population, and $CD49f^-$ HSCs as the lowest 30% $CD90^-$ $CD49f^-$ population. B) Representative example of the sorting strategy used to isolate LMPPs and MLPs progenitors at day 0 from $CD34^+$ CB cells.

3.3.3 Flow cytometry analysis

Flow cytometry analyses were performed using the BD LSR II Analyser with the BD LSR II HTC Analyser (BD Biosciences) in case of high-throughput analyses. The BD LSR Fortessa available at NIHR Cambridge BRC Cell Phenotyping Hub facility was also used occasionally. Unstained cells and compensation beads (BD Biosciences) were used for compensation and as controls to set appropriate gates. Data were analysed with FlowJo software (v 9.9 or v10). Where large number of samples were analysed, FlowJo data were exported and further analysed using R.

3.4 Single cell *in-vitro* functional assays

3.4.1 Single cell differentiation assays

Single cells were sorted into 96-well round-bottom plates (when monitoring the time of first division) or flat-bottom plates in 100 µl/well MEM (Myeloid (My)-Erythroid (Ery)-Megakaryocytic (Meg)) medium in order to support differentiation toward My-Ery-Meg lineages. MEM cytokine medium was prepared as follows: StemPro medium with nutrients supplement (Life Technologies) supplemented with cytokines (SCF 100 ng/ml, Flt3-L 20 ng/ml, TPO 100 ng/ml, IL-6 50 ng/ml, IL-3 10 ng/ml, IL-11 50 ng/ml, GM-CSF 20 ng/ml, IL-2 10 ng/ml, IL-7 20 ng/ml (all Miltenyi Biotec), EPO 3 units/ml (Eprex, Janssen-Cilag), h-LDL 50 ng/ml (Stem Cell Technologies), 1% L-Glutamine (Life Technologies) and 1% Pen/Strep (Life Technologies). After FACS sorting, plates were centrifuged for 5 minutes at 400 g and incubated at 37˚C.

To monitor the time of first division, single cells were visualised and counted manually every 8-12 h over 4 days using an inverted microscope. The time of first division was recorded for each cell and all the empty wells were excluded from the experiment. If the time of first division was initially recorded, cells were then transferred to a 96-wells flat bottom plate at day 4 after division kinetics analysis with addition of 50 µl of medium and cultured for a total of 3 weeks at 37˚C.

S1_1, S2_1, Diff1, S2_2, Diff2 populations were derived from bulk $CD49f^+$ Subset1 and $CD49f^+$ Subset2 cultured for 5 days in reduced cytokine MEM medium (lacking IL-11, IL-2 and IL-7). Single $CD49f^+$ HSCs, $CD49f^+$ Subset 1, $CD49f^+$ Subset2, S1_1, S2_1, Diff1, S2_2, Diff2 cells were then sorted into 96-well plates in 100 µl/well MEM cytokine medium and cultured for 3 weeks at 37˚C.

The type (lineage determination) and the size of the colonies formed were assessed after three weeks culture by high-throughput flow-cytometry using the BD LSR II HTC Analyser. All single cell derived colonies were harvested into 96 u-bottom plates. Each colony was stained using the antibody panel shown in Table 5. Plates were incubated for 20 minutes in the dark at RT with 50 µl/well of antibody mix and then washed with 100 µl/well of PBS + 3% FBS.

Lineage	Antibody (clone)	Dilution	Fluorochrome
Megakaryocytic	CD41(HIP8)	1:1000	FITC
Erythroid	GlyA (HIR2) *	1:1000	PE
hHSC marker	CD45 (HI30)	1:300	PE-Cy5
Monocytic	CD14 (M5E2)	1:1000	PE-Cy7
NK cells	CD56 (HCD56)	1:200	APC
Myeloid	CD11b (ICRF44)	1:300	APC-Cy7
Granulocytic	CD15 (MC-480)	1:200	BV421

Table 5: Antibodies used to assess the differentiation output in the MEM differentiation assay.

For each cell surface marker directed antibody, clone, dilution and fluorochrome are specified. All the antibodies listed were purchased from BioLegend except those indicated by * that were purchased from BD bioscience. Before use all the antibodies were titrated using appropriate control samples from human CB.

3.4.2 Single cell differentiation assay analysis

The type (lineage composition) and the size of the colonies formed were assessed by high-throughput flow-cytometry. The colony output was determined using the gating strategy shown in Figure 9. A single cell was defined as giving rise to a colony if the sum of cells detected in the $CD45^+$ and $GlyA^+$ gates was ≥30 cells. Ery colonies were identified as $CD45^-$ $GlyA^+$≥30 cells, Meg colonies as $CD41^+$≥30 cells, My colonies as [($CD45^+$ $CD14^+$) + ($CD45^+$ $CD15^+$)] ≥30 cells, NK colonies as $CD45^+$ $CD56^+$≥30 cells. My colonies were further classified as follows: Granulocytes (Gran) colonies were identified as $CD45^+$ $CD15^+$≥30 cells and $CD45^+$ $CD14^+$≤ 30, Monocyte (Mono) colonies as $CD45^+$ $CD15^+$≤30 cells and $CD45^+$ $CD14^+$≥30 and Monocyte/Granuocyte (MonoGran) as $CD45^+$ $CD15^+$≥30 cells and $CD45^+$ $CD14^+$≥30 cells. In the B cell assay, B cells were identified as $CD45^+$ $CD19^{++}$≥30 cells, My colonies as $CD45^+$ $CD11b^+$≥30 cells and NK colonies as $CD45^+$ $CD56^+$≥30 cells. All high-throughput screening flow cytometry data was recorded in a blinded way, and at analysis correlation between the colony phenotype and originating population was only performed at the final stage.

The analysis was done using FlowJo (v9.9) software. All the data obtained from the analysis were exported onto an excel file and when appropriate, correlated with the first time of division. The analysis was done using the R Studio software.

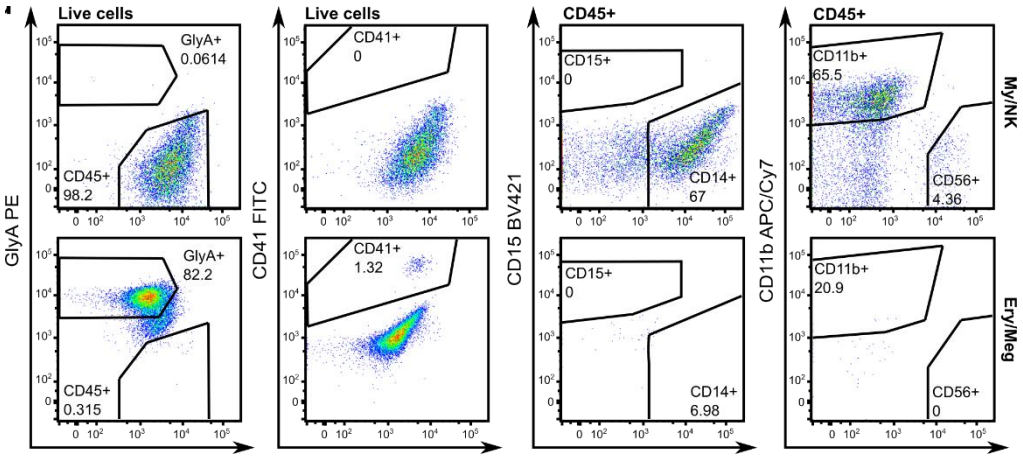


Figure 9: Gating strategy used for lineage determination in the MEM assay.

Representative examples of the gating strategies used to determine the lineages of colonies derived from $CD49f^+$ HSC single cells is shown. Erythroid colonies were identified as $GlyA^+$≥30 cells, megakaryocytic colonies as $CD41^+$≥30 cells, myeloid colonies as [($CD45^+$ $CD14^+$) + ($CD45^+$ $CD15^+$)]≥30 cells, NK colonies as $CD45^+$ $CD56^+$≥30 cells. Two representative examples (one My/NK and one Ery/Meg colony) are shown.

3.5 Cell size measurement

In order to measure the cell size of phenotypic LT-HSCs and ST-HSCs, 100 cells/well were sorted into a 384 well-plate directly into 50 µl of MEM cytokine medium prepared as described in paragraph 3.4.1 in presence or absence of Palbociclib 200 nM (PD0332991, Sigma).
Cell images were recorded every 24 hours in bright field (BF) with a 20x magnification on a Leica DMI300 B microscope using the MetaMorph Microscopy Automation and Image Analysis Software. The cell size was then measured on 25 cells/well using ImageJ and expressed as cell diameter size (µm).

3.6 Mitochondrial mass and mitochondrial activity measurements

Phenotypic LT-HSCs and ST-HSCs were sorted as previously shown in Figure 8.
As different dyes for mitochondrial mass (MitoTracker green FM) and mitochondrial activity (TMRM) are detected in the FITC or in the PE channel, the CD45RA antibody in the sorting panel C (see Table 4) was used either in the PE or in the FITC channel in order to leave one of these channels free for dye detection at 0 h after the sort.
The cells were then stained in 100 µl and analysed for the two above mentioned parameters on day 0 right after the sort, at 24 h and after 48 h in culture with MEM cytokine medium in presence and absence of Palbociclib 200 nM (PD0332991, Sigma).
The dyes and the adapted staining protocol used to determine the mitochondrial mass and the mitochondrial activity are shown in Table 6. After the staining the cells were washed with 1 ml of PBS + 3% FBS, re-suspended in 500 µl of PBS + 3% FBS and analysed by flow cytometry.

Dye	Concentration	Staining	Manufacturer	Channel
MitoTracker Green FM	20 [nM]	40 min 37˚C in medium	Life Technologies	FITC
TMRM	100 [nM]	40 min 37˚C in medium	Life Technologies	PE

Table 6: Dyes used to assess mitochondrial mass (MitoTracker Green FM), mitochondrial activity (TMRM).
The dyes used are listed in the table. For each one final concentration, staining conditions, manufacturer and channel are specified.

3.7 Index sorting analysis

Correlation analyses between index sorting data, division time and differentiation assays were performed with R (version 3.3.2). Cell-surface marker levels information from the indexed cells was recorded in *.csv* files during FACS sorting. Data collected from different sorts and on different days were normalised using a normalisation pipeline established in (Nestorowa et al. 2016): data obtained from each experiment were logicle transformed with the *estimateLogicle* function (m=5.1) of the FlowCore package and then then normalised using the ComBat function from the sva package. Principal

Component Analysis (PCA) was performed with the *prcomp* function on the normalised data, and vector loadings were obtained from the *prcomp* rotation variable. Dimensionality reduction analysis to produce a diffusion map was performed in the *bglab* package following methods described by (Haghverdi, Buettner, and Theis 2015) using a local sigma informed by the 5 nearest neighbours.

3.8 Mice

NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ (NSG) mice were purchased from Charles River or bred and maintained at the University of Cambridge animal facility. Experimental cohorts consisted of age-matched female mice of 12-16 weeks of age at the time of transplantation. All mice were housed in a Specific-Pathogen-Free (SPF) animal facility and experiments were conducted under UK Home Office regulations. This research has been regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB).

3.9 Xenograft transplantation assays

3.9.1 CD49f$^+$ Subset1 and CD49f$^+$ Subset2 transplantation assays

To assess human CD49f$^+$ engraftment *in-vivo*, NSG mice were sublethally irradiated (2.4 Gy). After 24 h from irradiation mice were anesthetised with isoflurane and injected intrafemorally (IF) with the indicated doses of the specified populations and injected subcutaneously with 0.1 mg/kg buprenorphine (Animalcare). As all mice shared the same genetic background and were age-matched females, no specific randomisation nor blinding was performed. For Limiting Dilution Analysis (LDA) experiments, all sorted subsets where maintained overnight in X-VIVO 10 medium (Lonza) supplemented with 1% BSA (Roche), 1% L-glutamine (Life Technologies), 1% Pen/Strep (Life Technologies), SCF 100 ng/ml, Flt3-L 100 ng/ml, TPO 50 ng/ml and IL-7 10 ng/ml before injection.

Mice transplanted with saturating doses of CD34$^+$ HSCs were given intraperitoneal injections of either PBS or EPO (20 units/injection, Eprex, Janssen-Cilag) every other day for 2 weeks prior sacrifice. For LDA experiments, mice analysed at 20 weeks post-transplantation received 8 intraperitoneal EPO injections (20 units/injection) in the 4 weeks prior to sacrifice. At 20 weeks post transplantation the injected femur was harvested, and the BM was flushed out.

3.9.2 Cultured LT-HSCs transplantation assays

To assess human cultured LT-HSC engraftment *in-vivo*, NSG mice were sublethally irradiated (2.4 Gy). After 24 h from irradiation mice were anesthetised with isoflurane and injected intrafemorally (IF) with the indicated doses of the specified populations and injected subcutaneously with 0.1 mg/kg buprenorphine (Animalcare). As all mice shared the same genetic background and were age-matched females, no specific randomisation nor blinding was performed. LT-HSCs were sorted in bulk and

cultured for 24 or 72 h in MEM media (section 3.4.1) with or without PD0332991 until the injection day. At 18 weeks post transplantation the injected femur was harvested, and the BM was flushed out.

3.9.3 Secondary transplants

Secondary transplantation was performed by purifying $CD34^+$ cells from individual primary animals using the Micro Beads $CD34^+$ selection kit (Miltenyi Biotec), then $CD34^+CD38^-$ cells were sorted by flow cytometry from each primary mouse. All the cells sorted from each mouse (1000-30000 cells) were IF injected into secondary recipients. Animals were then sacrificed 12 weeks post-secondary transplantation.

3.9.4 In-vivo engraftment analysis

Cells obtained from BM and injected bone (IB) from all assays were stained using the panel shown in Table 7.

Lineage	Antibody (clone)	Dilution	Fluorochrome
B cells	CD19 (HIB19)	1:300	Alexa 700
B cells	CD19 (HIB19)	1:200	FITC
Erythroid	GlyA (HIR2) *	1:1000	PE
hHSC marker	CD45 (HI30)	1:300	PE-Cy5
hHSC marker	CD45 (HI30)	1:500	BV510
Monocytic	CD14 (M5E2)	1:1000	PE-Cy7
T cells	CD3 (HIT3a)	1:100	APC-Cy7
Myeloid`	CD33 (P67.6) *	1:200	APC

Table 7: Antibodies used to assess the engraftment of transplanted mice.

For each cell surface marker directed antibody, clone, dilution and fluorochrome are specified. All the antibodies listed were purchased from BioLegend except those indicated by * that were purchased from BD bioscience. Before use all the antibodies were titrated using appropriate control samples from human CB.

To ensure appropriate detection of low levels of human engraftment, two distinct antibodies against the human HSC marker CD45 were used. Cells were considered of human origin if positive for both ($CD45^{++}$). Mice were considered engrafted if ($\%CD45^{++} + \%GlyA^+$) ≥ 0.01 % and at least 30 cells were recorded in these gates. Mice used as controls and not injected were sometimes included and defined as irradiated but non-transplanted mice. BM staining of control mice was sometimes used to define staining background. None of these mice met the criteria for engraftment.

To determine the lineage composition of the human graft, the following parameters were adopted: My lineage: $CD45^{++}$ $CD33^+ \geq 30$ cells; Ly lineage: $CD45^{++}$ $CD19^{++} \geq 30$ cells (positive for 2 distinct CD19 antibodies); Ery lineage: $CD45^-$ $GlyA^+ \geq 30$ cells or $CD45^-$ $CD71^+$ $GlyA^+ \geq 30$ cells.

3.10 Bulk RNA-sequencing analysis

Bulk RNA-seq expression data from LT- and ST-HSCs at 0 h (q), 72 h (a) and 72 h with 200nM PD0332991 (pq) were obtained from John Dick's Lab (Princess Margaret Cancer Centre, University Health Network, University of Toronto). These data were generated from 500 cells. Library preparation was done following the Smart-Seq2 protocol (Picelli et al. 2013) and using the low input Nextera Kit (Illumina). Differential expression between q vs a, q vs pq from both LT- and ST-HSCs and aLT-HSCs vs aST-HSCs populations was performed with DESeq v1.6.1, using a read counts cut-off of 25.
Venn diagrams were created to show the number of genes differentially expressed that are in common between different comparisons. Functional annotation was performed with DAVID (v 6.8) .
To measure how similar the activation process is at the transcriptional level between LT- and ST-HSCs, the fold-change in expression for each gene between the quiescent and activated state was calculated and correlation plots were drawn. The correlation coefficient was calculated with the *cor.test* function (two-sided Pearson method).

3.10.1 Short Time-series Expression profile analysis

Temporal expression profile analysis of bulk RNA-seq gene expression data was performed using the Short Time-series Expression Miner version 1.3.12 (Jason Ernst and Bar-Joseph 2006) on 13121 genes. This software uses a clustering method (J. Ernst, Nau, and Bar-Joseph 2005) that can differentiate between real and random patterns and clusters genes by assigning them to a series of pre-defined expression profiles. Similar profiles are grouped together to form clusters and biological significance of the genes associated with the cluster is assessed using a Gene Ontology (GO) enrichment analysis.
If the number of genes assigned to a profile exceeds the number of genes that are expected to occur by chance, the profile is significant. Analysis was done using STEM default parameters. The significance level for profile assignment was set to be 0.05.

3.10.2 Gene Set Enrichment Analysis (GSEA) on bulk RNA-seq data

Gene Set Enrichment Analysis (GSEA) allows to determine if a set of genes shows statistically significant differences between two biological states. Raw expression files were used and the analysis was run with either the C2 curated gene sets sub collection of canonical pathways or the C3 transcription factor motif gene sets from the MSigDB60 (Subramanian et al. 2005). Significantly enriched gene sets at FDR< 25% were taken into consideration for biological interpretation.

3.11 ScRNA-sequencing library preparation

The protocol used to prepare libraries for scRNA-seq was adapted from the Smart-seq2 protocol of (Picelli et al. 2013). Single cells were sorted into 96 well PCR plates in a configuration determined to minimise batch effects due to the position of the cells on the plate. Single cells were sorted into 4 µl of

lysis buffer/well prepared containing 0.4% (v/v) Triton X-100, 2 U/µl RNase inhibitor (Clontech), 5 mM DTT, 1 mM dNTP and stored at −80 °C. ERCC used as external RNA controls (Ambion, Life Technologies) were diluted to a final concentration of 1:3,000,000. The PCR purification step after cDNA amplification was done with 20 µl of Ampure XP beads (ratio 1: 0.6/0.7, Beckman Coulter). The success of cDNA preparation was confirmed by optimal cDNA signal detected by a 2100 Bioanalyzer with High-sensitivity DNA chip (Agilent). cDNA concentration for each single-cell was determined using the Picogreen quantification kit (ThermoFisher). An optimal range of 0.07-0.120 ng/µl was considered good to proceed with library preparation. Dilution plates were calculated and prepared in order to have most of the cells in the plate within the ideal cDNA concentration range needed for library preparation. Illumina library preparation was carried out following the Illumina Nextera XT DNA sample preparation protocol. Library size distribution was checked on an Agilent high-sensitivity DNA chip and the concentration of the indexed library was determined using the KAPA library quantification kit (Kapa Biosystems). The sequencing was done using the Illumina Hiseq 4000 system and performed by CRUK Cambridge Institute Genomics Core.

3.12 Single cell RNA-sequencing analysis

3.12.1 Single cell RNA-sequencing analysis for CD49f$^+$ HSCs, CD49f$^+$ Subset1, CD49f$^+$ Subset2, Subset2, MLPs and LMPPs

NOTE: Bioinformatic analysis was performed by Elisa Laurenti and Kendig Yen Chi Sham.

Reads alignment was performed using GSNAP (T. D. Wu and Nacu 2010), and read counts were generated with HTseq (S. Anders, Pyl, and Huber 2015). All analyses below were performed in R (version 3.3.2) with the *bglab* package (https://github.com/wjawaid/bglab). Blinded quality control (QC) was performed on the single cell/sample group and single cells/samples were retained for further analysis if $> 2x10^5$ reads mapped to a gene feature, there were > 20% of genes over the total number of reads and < 20% of mitochondrial genes over mitochondrial + nuclear genes. Data were normalised for sequencing depth using size factor calculated on endogenous genes (Simon Anders and Huber 2010). Batch effects were corrected using the *ComBat* function from the *sva* package on all genes that had >1 count. Highly variable genes (HVGs) were selected fitting a GAM model assuming a quadratic relationship between log coefficient of variance (CV) and log mean expression for ERCC spike-in genes, with the function *techVar (bglab* package) setting the MeanForFit parameter at 10 (Brennecke et al. 2013). All dimensionality reduction techniques were applied to HVGs. Differential expression between the 2 populations was performed with DESeq2 using the *doDESeq* wrapper function of the *bglab* package (Love, Huber, and Anders 2014). Genes were considered differentially expressed if FDR < 0.05. Expression of selected genes was represented using violin plots (*vioplot* package) of the log10 (1 + normalised_counts).

ICGS analysis was run with AltAnalyze software (http://altanalyze.readthedocs.io/en/latest/ICGS/) using normalised and batch corrected counts as input and default parameters. Cell cycle genes were excluded using the most stringent parameter. For the analysis including CD49f$^+$ Subset1 and CD49f$^+$ Subset2 single cells, the software identified 6 minor clusters that were regrouped into 3 major clusters. For GSEA, genes were ranked by the DESeq2 statistic and pre-ranked GSEA was run using either the C2 curated gene sets sub collection of canonical pathways or the C3 motif gene sets from the MSigDB60 (Subramanian et al. 2005) population specific signatures (Laurenti et al. 2013, 2015) or lineage-priming module (Velten et al. 2017).

3.12.2 Single cell RNA-sequencing analysis for LT-HSC time course

For the generation of the time course scRNA-seq data two experiments were performed and sequenced separately (A and B below).

Alignment was performed using G-SNAP (T. D. Wu and Nacu 2010) and read counts were generated using HTSeq (S. Anders, Pyl, and Huber 2015) with Ensembl genes (Zerbino et al. 2018)

Quality control (QC) was performed independently for each experiment in R (version 3.6.0) (https://www.r-project.org) with *bglab* package ([https://github.com/wjawaid/bglab]). This was performed blinded of the single-cell/sample group and single cells were retained for further analysis if: $> 2 \times 10^5$ reads mapped to a gene feature, ratio of genes to total number of reads greater than 0.3 (A) or 0.2 (B), ratio of mitochondrial genes to mitochondrial and nuclear gene less than 0.15 (A) and 0.2 (B), number of nuclear genes to total mapped genes ratio greater than 0.75 (A) and 0 (B), number of genes with 10 reads per million greater than 2000 (A) and 0 (B), spike-ins to mapped reads ratio less than 0.2 (A) and 1 (B), nuclear genes to mapped reads ratio greater than 0 (A and B).

Exp A	**LT 0 h**	**LT 6h**	**LT 24h**	**LT 48h**	**LT 72h PD**	**LT 72h UNTR**
Before QC	72	72	72	72	72	72
After QC	29	49	28	20	27	42

Exp B	**LT 0h**	**LT 6h**	**LT 24h PD**	**LT 24h UNTR**	**LT 72h UNTR**
Before QC	95	95	95	95	95
After QC	57	87	80	58	82
Excluding additional outliers	56	85	80	58	82

Table 8: Number of cells analysed by scRNA-seq before and after quality control in all the time points and conditions analysed.
The number of cells sequenced and the number of cells that passed QC are shown separately for Experiment A (top) and B (bottom).

For data visualisation and pseudotime calculation, the two datasets were combined using the Scanpy module in python (Wolf, Angerer, and Theis 2018). Genes were filtered out if expressed in < 3 cells (scanpy method: *filter_genes*, parameters: min_cells = 3). Counts were logarithmized to the base *e* after adding a pseudo-count of 1 (scanpy method: *log1p*, default parameters). Batch correction was performed

with Scanpy function combat with parameters: covariates = timepoint. The original implementation of the algorithm in python can be found at https://github.com/brentp/combat.py. HVGs were selected using the Scanpy method: *highly_variable_genes*.

	Parameters for scanpy function: highly_variable_genes
Combat without PD	minimum_mean=0.05, maximum_mean=13, minimum_dispersion=0.1, maximum_dispersion=3
Combat with PD	minimum_mean=0.05, maximum_mean=6, minimum_dispersion=0.1, maximum_dispersion=3

Table 9: Table of the parameters used for the Scanpy function *highly variable genes.*

PCA was calculated for the combined datasets (Scanpy method: *pca*, parameters: SVD solver='arpack'). Nearest neighbours for each cell were calculated with the python package *bbknn* with parameters: approx= False, metric='euclidean' (Park et al. 2018). Diffusion map was calculated in 3 dimensions (Scanpy method: *diffmap*). The *dpt* function in Scanpy was used to obtain the pseudotime index for each cell (parameters: n_dcs= 3 (to match diffusion map dimension)). UMAP visualisation was performed using the function *sc.tl.umap* (data_neighbor, n_components=3).

For differential expression analysis between different time points, raw counts generated from high-throughput sequencing (HTSeq) were used for filtering out genes that were expressed in < 3 cells. Differential expression results were generated by DESeq2 (Love, Huber, and Anders 2014) and where applicable, the batch effect was accounted for.

Analysis	**Design matrix input for DESeq2**
no PD	design = ~ batch + timepoint
Exp A	design = ~ timepoint
Exp B	design = ~ timepoint

Table 10: Design matrix input for DESeq is shown for each analysis performed.

Cell cycle phase assignment was done using *scanpy.api.tl.score_genes_cell_cycle* function, implementing an approach used by Seurat package (Satija et al. 2015). The list used for the function can be found here: https://github.com/theislab/scanpy_usage/blob/master/180209_cell_cycle/data/regev_lab_cell_cycle_genes.txt. The list contains a total number of 97 genes, compiled as per (Tirosh et al. 2016).

The DESeq2 object from the differential analysis above without 48 h and PD timepoints was used for gene expression pattern analysis. The function *vst* was applied to normalise gene counts with parameter: blind=FALSE. All differentially expressed genes between any pair of timepoints were extracted from the DESeq2 object. The list was then filtered for adjusted p-value < 0.05. The final list contained 10010 unique genes. The *degPatterns* function in the R package *DEGreport* (Pantano 2019) was called with parameters: time='timepoint', eachStep=TRUE and 11 gene expression clusters were found.

3.12.3 ClueGO analysis

Non-redundant biological terms analysis on differentially expressed genes within each pattern of expression identified using the DEGreport R package (see section 3.12.2) was performed with the Cytoscape plug-in, ClueGO (v 2.5.2) (Bindea et al. 2009). P-value < 0.1 was used to select significant pathways. The analysis was run against KEGG, REACTOME, Biological processes, Immune System Processes and Molecular Function Terms ontologies. The GO Term Fusion option was used.

3.13 Statistical analysis

Statistical analysis was performed with R or Graph Pad Prism, after verification that statistical tests were appropriate given the distribution of the data and variance between groups.

4 RESULTS

NOTE: in this section results will be presented in two separate chapters. Chapter A will address results addressing Aim A and Chapter B the results addressing Aim B.

Chapter A: The first myelo-lymphoid lineage restriction already occurs in the human haematopoietic stem cell compartment

NOTE: Sections of this chapter report data published in Belluschi et al., Nat. Communications, 2018. The work related to this publication was done in collaboration with Emily Calderbank, PhD student in Elisa Laurenti group. While I mainly worked on the characterisation of functional and molecular heterogeneity of the CD49f$^+$ HSC compartment, similar experiments to characterise the HSC/MPP pool were done by Emily Calderbank. Here, I will present results obtained from the analysis of the CD49f$^+$ HSC compartment, however, to better understand the results of the work published, references to data obtained from the analysis of the HSC/MPP pool will be presented too.

4.1 Functional and transcriptional characterisation of the CD49f$^+$ HSC compartment at the single cell level

4.1.1 My/Ly and My/Ery in-vitro *differentiation is determined by pre-existing heterogeneity within the CD49f$^+$ HSC compartment*

The purest human HSC compartment known to date consists of phenotypic LT-HSCs and it is also referred to as the CD49f$^+$ HSC compartment. Evidence in the field showed heterogeneity and variability at the single cell level within the human HSC/MPP pool and the CD49f$^+$ HSC compartment in relation to cell division properties, cell surface markers expression and cell fate decisions (Notta et al. 2016; Velten et al. 2017) (see section 1.2). This suggests a possible link between these properties. However, to date, no one has comprehensively characterised these two compartments in terms of *in-vitro* differentiation potential and division kinetics at the single cell level. As published data from Laurenti et al. have shown, the kinetics of first division are differentially regulated within the HSC/MPP pool and specifically between CD49f$^+$ HSCs and CD49f$^-$ HSCs (Figure 3). Here I present a comprehensive analysis of the differentiation potential and time to first division of single CD49f$^+$ HSCs.

To characterise the *in-vitro* differentiation potential of single CD49f$^+$ HSCs (CD19$^-$ CD34$^+$ CD38$^-$ CD45RA$^-$ CD90$^+$ CD49f$^+$) towards myeloid (My), lymphoid (Ly), megakaryocytic (Meg) and erythroid (Ery) lineage I optimised and repurposed an assay developed for the assessment of My, Ery and Meg potential (Notta et al. 2016). Moreover, to detect variability in division kinetics, the time of first division for each single cell was recorded before assessing the differentiation potential through *in-vitro* functional assays (Figure 10 A). This is important to identify possible correlations between cell cycle progression

and cell fate decisions that could be present in HSCs, similar to what has been previously demonstrated in other stem cells (Pauklin and Vallier 2013; Dalton et al. 2015).

819 $CD49f^+$ HSCs from 4 independent $CD34^+$ human CBs samples were index-sorted and for each cell the time to first division, the differentiated colony output and the cell surface intensity of 8 proteins (CD19, CD34, CD38, CD45RA, CD90, CD49f, CD10, CLEC9A) was recorded. CLEC9A, a surface receptor for which the mRNA expression was found to be significantly higher in $CD49f^+$ HSCs ($CD19^- CD34^+ CD38^- CD45RA^- CD90^+ CD49f^+$) than in $CD49f^-$ HSCs ($CD19^- CD34^+ CD38^- CD45RA^- CD90^- CD49f^-$ (Figure 10 B) was included in this panel. The expression of CLEC9A, a SYK-coupled C-type lectin receptor, was only reported within the dendritic cell compartment to date. Flow cytometry analysis confirmed that approximately 80% of $CD49f^+$ HSCs are $CLEC9A^+$ compared to 50% of $CD49f^-$ HSCs while in progenitor populations showed an even lower expression of this marker (Figure 10 C). Single cells were cultured in liquid medium supporting the growth of My, Ery, Meg and NK cells and the colony output was assessed after 3 weeks of culture via high throughput screening flow cytometry (markers to distinguish differentiated cell types are listed in Table 5). Approximately 80% of single cultured $CD49f^+$ HSCs were able to produce a colony, indicating that the remaining 20 % of the cells produced very small colonies ([$CD45^+$ and $GlyA^+$] <30 cells, see Methods) that were not considered in the analysis or died due to FACS related stress before dividing and differentiating (Figure 10 D). Analysis of the colonies produced showed that less than 1% of $CD49f^+$ HSCs can originate quadrilineage colonies and around 10% of multilineage colonies were of the My/NK/Ery type while My/NK and My/Ery colonies were present in similar proportions and My/Ery/Meg accounted for around 20% of the multilineage colonies. The only unilineage colonies produced in this assay were My-only (Figure 10 E). Interestingly, Ery and NK cells were always found together with My cells, whereas Meg differentiation, less efficient than all other lineages, was almost always associated with Ery output. I also included markers to distinguish monocytic (Mon), granulocytic (Gran) or mixed MonGran colonies: more than 80% of the My colonies were represented by MonGran (Figure 10 F).

Interestingly, the combined analysis of the time to first division and the colony output revealed that single $CD49f^+$ HSCs generating colonies containing Ery cells complete their first division significantly later than those producing colonies containing NK cells ($p<0.001$, Figure 10 G, H). Despite the marked delay in the time to first division of cells differentiating towards the Ery lineage, colonies of the My/Ery type were significantly larger than any other type of colony (Figure 10 I) suggesting that a maximal expansion is present along the Ery differentiation branch.

Overall, in this assay, single $CD49f^+$ HSCs show a reproducible pattern of colony formation which is indicative of pre-existing heterogeneity in the differentiation output of cells within this compartment. This variability not only extends to the type of colony produced but also to the time of first division, suggesting a possible correlation between cell cycle regulation and fate choices in HSCs.

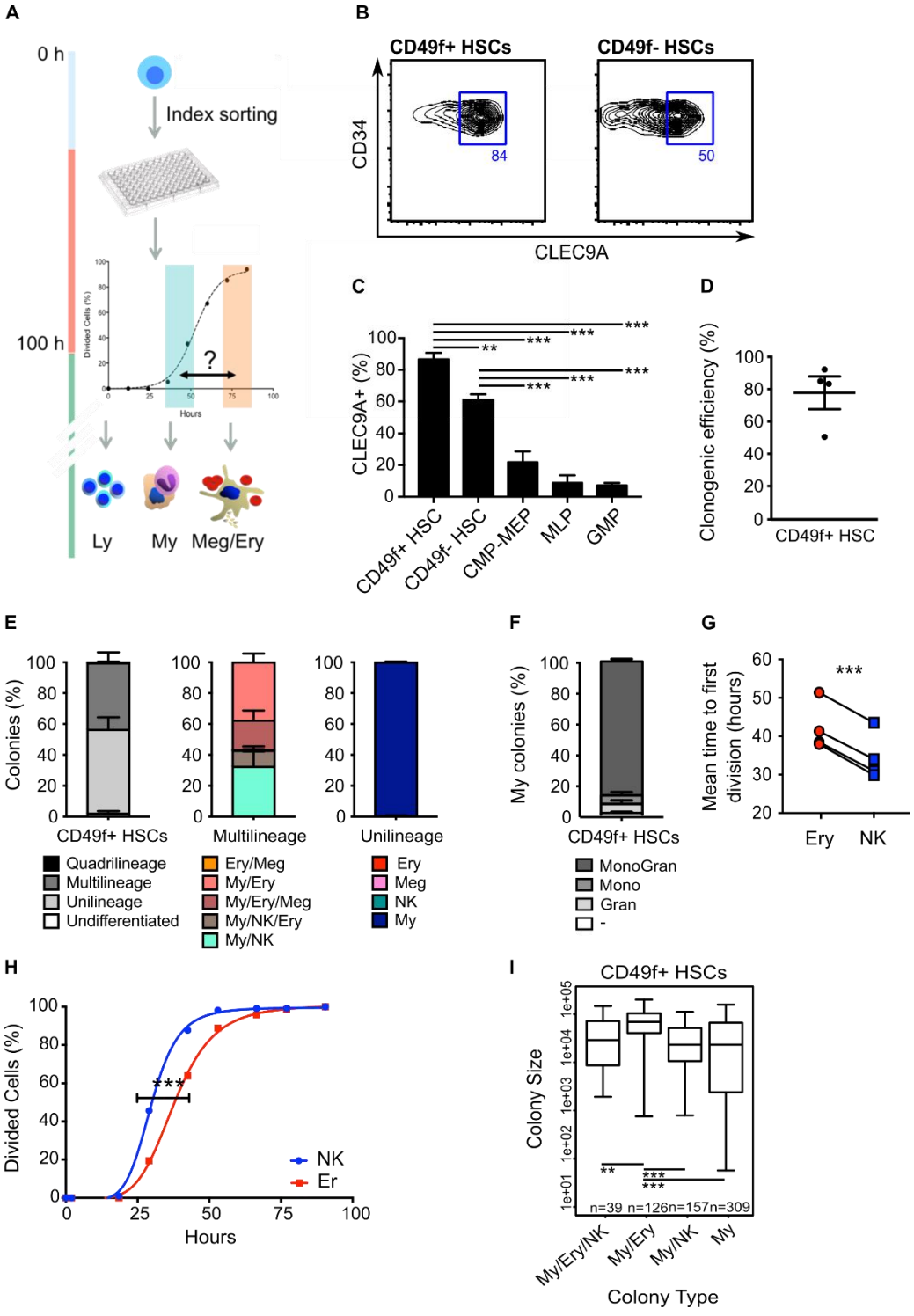


Figure 10: ***In-vitro*** **differentiation output characterisation of single CD49f$^+$ HSCs.**

A) Experimental design used to characterise the differentiation output of single CD49f$^+$ HSCs. Single cells were sorted at 0 h and the surface markers expression was recorded by index sorting. The time of first division for each single cell was monitored and recorded over 100 h in culture and the type of colony obtained was assessed by high-throughput flow cytometry after 3 weeks of culture. B) A representative example of CLEC9A cell surface expression in CB derived CD49f$^+$ HSCs (top left, n=307 cells) and CD49f$^-$ HSCs (top right, n=296 cells) is shown together with the percentage of CLEC9A$^+$ cells (blue gate). C) Percentage of CLEC9A$^+$ cells within the indicated populations. Mean ± SEM is shown. Statistical significance is calculated by one-way ANOVA with Tukey's multiple comparison (**p<0.01, ***p<0.001). D) *In-vitro* single cell differentiation assay clonogenic efficiency of single CD49f$^+$ HSCs (n=1104 cells from 4 independent CB samples). Mean ± SEM is shown. E) Percentage of colonies of the indicated type obtained from *in-vitro* differentiation of single CD49f$^+$ HSCs (n =819 colonies from 4

independent CB samples). Mean ± SEM is shown. F) Percentage of My colonies of the MonoGran, Mono, Gran and undetermined (-) type derived from single CD49f$^+$ HSCs (n=725 colonies from 4 independent CB samples). Mean ± SEM is shown. G) Mean time of first division of single CD49f$^+$ HSCs producing colonies containing cells from the indicated lineages (EC50 of non-linear fit of cumulative first division kinetics); n=4 experiments with independent CB samples (with respectively 19, 98, 34, 230 single cells per experiment). Statistical significance is calculated by two-sided paired t-test (***p < 0.001). H) Representative example of first division kinetics for single CD49f$^+$ HSCs producing colonies containing one of the two indicated lineages (Ery n=91, NK n=139). The graph is a non-linear fit cumulative curve. Statistical significance is calculated by extra sum-of-square F test (***p<0.001). I) Size of the different types of colonies derived from single CD49f$^+$ HSC from 4 different experiments and independent CB samples. Size is shown as number of cells per colony. Median, interquartile and 5-95 percentiles are shown. Statistical significance is calculated by Kruskal- Wallis test with multiple comparison (**p<0.01, ***p<0.001).

4.1.2 Pre-existing heterogeneity within the CD49f$^+$ HSC compartment correlates with cell surface expression of CD34 and CLEC9A

Evidence of a reproducible pattern of colony formation suggested pre-existing heterogeneity in lineage output within the CD49f$^+$ HSC compartment which could potentially correlate with the expression of specific cell surface markers. To investigate this, a principal component analysis (PCA) of the cell surface markers expression at the time of sort was performed. This showed that single CD49f$^+$ HSCs generating My-only colonies are evenly scattered across the PCA space whereas those producing My/Ery colonies and My/Ly colonies are unequally distributed along the two principal components PC1 and PC2 (Figure 11 A). Cells producing multilineage My/NK/Ery or My/NK/Ery/Meg colonies show a similar distribution as those producing My/Ery colonies (Figure 11 A). Overall, this indicates a functional polarisation of these cells within the PCA space with cells producing predominantly multilineage and My/Ery colonies on one side and cells producing My/Ly colonies on the other side. A similar functional polarisation was observed when single cells of the less pure HSC/MPP compartment were analysed similarly, indicating that this pre-existing level of heterogeneity extends also beyond the CD49f$^+$ HSC compartment. Although, because of the fast kinetics of division of the HSC/MPP compartment, no correlation between the time of first division and cell fate was identified in this compartment (data not shown).

Analysis of the PCA vector loadings showed that cell surface expression of two markers, CD34 and CLEC9A, drives most of the variance on PC1 and PC2 within CD49f$^+$ HSCs (Figure 11 C). A comparison of the cell surface expression of these two markers between CD49f$^+$ HSCs that generated Ery or My/Ery and or My/NK or NK colonies showed that the former had significant lower levels of CD34 and higher levels of CLEC9A (Figure 11 D). Moreover, milder but still significant shifts in the expression of these two markers were also present between cells that generated My/NK colonies with a Ly-bias (the number of My cells < number of Ly cells), where the number of NK cells in the colonies exceeds the number of My cells, and those with My-bias (My cells > number of Ly (NK) cells) but also between cells giving rise to My/Ery colonies with an Ery-bias, where the number of Ery cells in the colony exceeds the number of My cells, and those with a My-bias (Figure 11 E).

In summary, these results uncover a pre-determined functional polarisation within the CD49f$^+$ HSC compartment that separates single cells with My/Ly potential from those with My/Ery or multilineage

potential. Moreover, the lineage output seen in *in-vitro* functional assays correlates with different levels of expression of CLEC9A and CD34 surface markers, also indicating that single cells with distinct lineage outputs are distributed over a continuum of anticorrelating CLEC9A and CD34 cell surface expression.

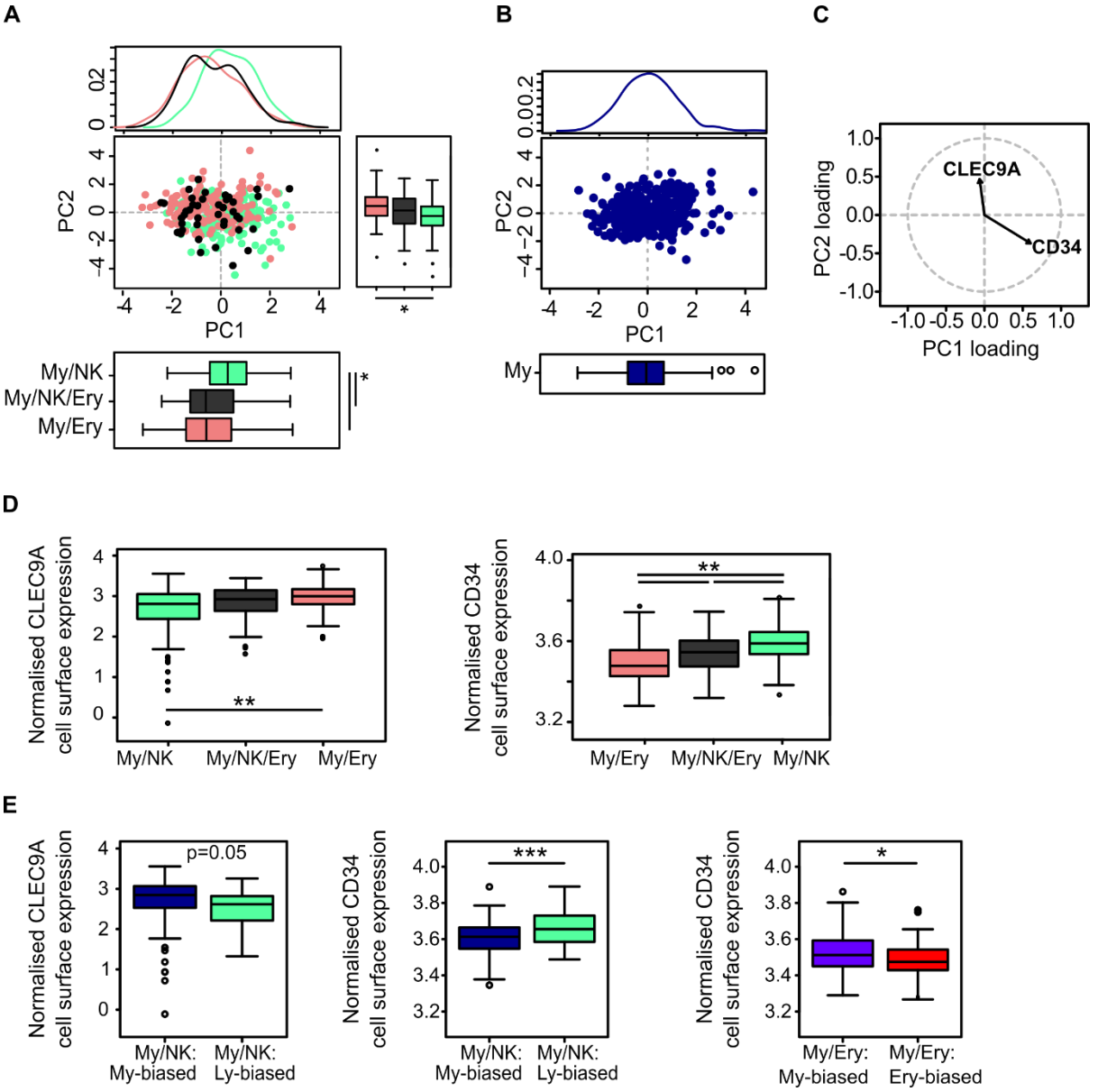


Figure 11: The continuum of differentiation outputs within the human CD49f⁺ HSCs compartment correlates with gradual changes in surface expression of CLEC9A and CD34 markers.

A-C) Principal component analysis (PCA) of the surface marker expression recorded at the time of sort of single CD49f⁺ HSCs (n=714 colonies from 4 independent CB samples). Each type of colony produced by a single cell after culture is indicated by different colours (here, My/NK/Ery includes My/NK/Ery and My/NK/Ery/Meg colonies). Top and bottom panels indicate respectively density plots and boxplots of PC1 and PC2 values from single cells producing the indicated types of colonies. Statistical significance is calculated by Kruskal–Wallis test with multiple comparisons (*p < 0.05). B) Single cells that produced My only colonies are showed here (n=309 colonies). C) The two cell surface markers with the highest PC loadings are shown here. D) Normalised intensity of CLEC9A (left) and CD34 (right) surface expression at the time of the sort for single CD49f⁺ HSCs producing the indicated colonies (n=714 colonies from 4 independent CB samples). The median fluorescence intensity shift between My/Ery and My/NK is 40% for CLEC9A and 13% for CD34. Statistical analysis is calculated by Kruskal-Wallis test

with multiple comparison (**p<0.01). E) Normalised intensity of CLEC9A (left) and CD34 (right) surface expression at the time of the sort for single CD49f$^+$ HSCs producing the indicated colonies. My/NK colonies are defined: as My-biased if the number of My cells > number of Ly (NK) cells within that colony; as Ly-biased if the number of My cells < number of Ly cells. Similarly, My/Ery colonies are defined as Ery-biased if the number of My cells < number of Ery cells (n=93 My/Ery colonies and n=170 My/Ly colonies from 4 independent CB samples). Median fluorescence intensity shifts between My/NK: My-biased and My/NK: Ly-biased is 25% for CLEC9A and 4% for CD34; between My/Ery: My-biased and My/Ery: Ery-biased is 4% for CD34. Median, interquartile and 5-95 percentiles are shown. Statistical significance is calculated by two-sided unpaired t-test (*p<0.05, ***p<0.001).

4.1.3 Anticorrelating cell surface expression of CLEC9A and CD34 defines a continuum of lympho-myeloid priming within the transcriptional space of CD49f$^+$ HSCs

After identifying pre-existing heterogeneity within the CD49f$^+$ HSC compartment and its correlation with CLEC9A and CD34 surface expression, I performed scRNA-seq to investigate if these observations would also be mirrored at the transcriptional level.

I index-sorted a total of 192 single CD49f$^+$ HSCs and performed scRNA-seq using an adapted version of the SmartSeq2 protocol (Picelli et al. 2013). Quality control (QC) analysis showed that 88% of the cells passed QC (see methods). To compare the results seen at the functional level with the transcriptomic data obtained from the analysis of single CD49f$^+$ HSCs, three different dimensionality reduction techniques were used (tSNE, PCA and diffusion map). The surface expression of CD34 and CLEC9A, as recorded by index sorting, was overlaid on the cells represented in the transcriptional space. The transcriptional space of CD49f$^+$ HSCs is a continuum where CLEC9A levels progressively decrease from one extreme to the other while CD34 levels progressively increase in the opposite direction to CLEC9A (Figure 12 A-F). This is in concordance with previous results obtained from the *in-vitro* functional assays. Iterative Clustering and Guide-gene Selection (ICGS) algorithm (Olsson et al. 2016) was then applied to identify, in an unbiased manner, possible clusters of cells driven by expression of specific gene signatures within this dataset. From this analysis two clusters comprising 53 and 116 cells were identified and classified respectively as Cluster I and Cluster II (Figure 12 H). The surface expression of CLEC9A and CD34 was compared between cells of each cluster. Cell within Cluster I showed significantly higher levels of CLEC9A surface expression but lower levels of CD34 than cells within Cluster II (Figure 12 G).

In conclusion, the transcriptomic analysis of CD49f$^+$ HSCs showed that the transcriptional landscape of these cells is represented as a continuum of transcriptional states which is polarised at its extremes. Here, opposing gradients of CLEC9A and CD34 cell surface expression correlate with differences in genes expression and the observed polarisation of the transcriptional landscape matches the one previously identified from single cell *in-vitro* differentiation assays in CD49f$^+$ HSCs.

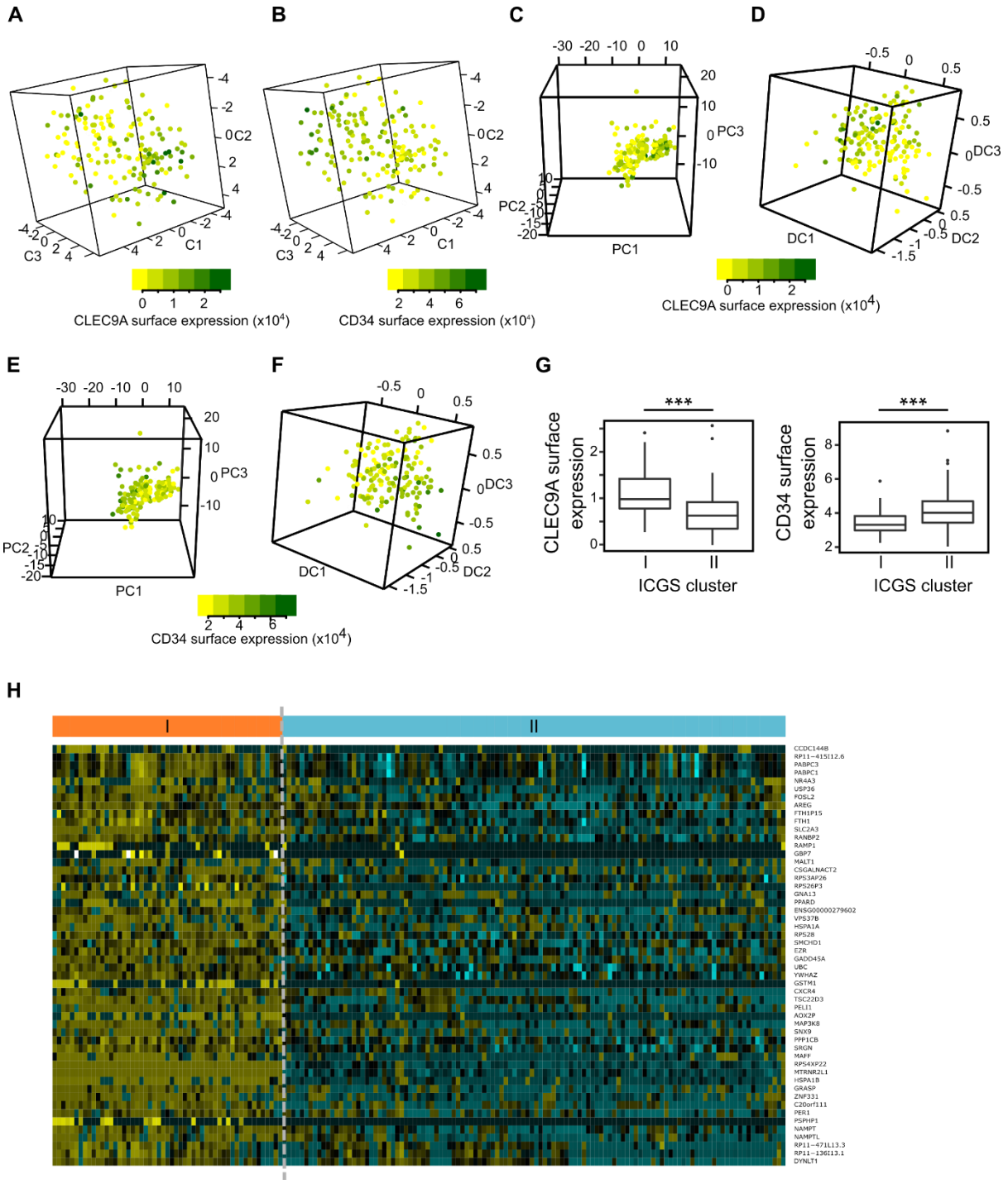


Figure 12: Anticorrelating CLEC9A and CD34 surface expression polarises the transcriptional space of single CD49f$^+$ HSCs.

A, B) tSNE 3D representation of the transcriptional landscape of CD49f$^+$ HSCs. CLEC9A (A) and CD34 (B) surface expression levels at the time of sort are overlaid on the tSNE plot (n= 169 cells). C, D) CLEC9A cell surface marker expression overlaid on 3D PCA (C) or diffusion map (D) representation of CD49f$^+$ HSCs. E, F) CD34 cell surface marker expression overlaid on 3D PCA (E) or diffusion map (F) representation of CD49f$^+$ HSCs. G) Cell surface expression levels of CLEC9A (left) and CD34 (right) in the two indicated ICGS clusters of single CD49f$^+$ HSCs. Median, interquartile and 5–95 percentiles are shown. Statistical significance is calculated by two-sided unpaired t-test (***$p < 0.001$). H) Expression heatmap of guide genes selected by the ICGS algorithm performed on CD49f$^+$ HSCs (n=169 cells). Columns represent single cells and rows represent genes. The top panel bar shows the two clusters identified by ICGS (Cluster I and Cluster II).

4.2 Subsets of CD49f$^+$ HSCs with My/Ly and My/Ery differentiation *in-vitro* can be prospectively purified based on the surface expression of CLEC9A and CD34

A bilineage output *in-vitro* could indicate that the single cell of origin is either a multipotent HSC with a strong lineage bias or also a lineage restricted HSC. To discriminate between these hypotheses, I looked into developing new prospective purification strategies to isolate subset populations of the CD49f$^+$ HSC compartment with specific behaviours *in-vitro.*

The correlation between the surface expression of CLEC9A and CD34 and the type of colony produced prompted me to search for possible flow cytometry gating strategies based on the expression of these two markers that would enrich for subsets of CD49f$^+$ HSCs with preferential differentiation towards My/Ly and My/Ery. Using the single cell data obtained from previously shown *in-vitro* functional assays (Figure 10), all possible gating thresholds along the CLEC9A/CD34 spectrum were analysed *in-silico* to identify the sorting strategies that would maximise the enrichment for bipotent lineage output *in-vitro* (My/Ery or My/NK) (analysis done by Blanca Pijuan-Sala). From this analysis, I chose two gating strategies for a CD34^{lo} CLEC9A^{hi} fraction (hereafter referred to as CD49f$^+$ Subset1) and a CD34^{hi} CLEC9A^{lo} fraction (hereafter referred to as CD49f$^+$ Subset2). These were predicted to yield the highest enrichment respectively for single cells with My/Ery and My/Ly potential *in-vitro.* Similar populations were also identified within the HSC/MPP compartment and they were named Subset1 (CD34^{lo} CLEC9A^{hi} fraction) and Subset2 (CD34^{hi} CLEC9A^{lo} fraction) (Figure 13 A). Intermediate fractions between these two gates were predicted to contain equal proportions of HSCs with My/Ly and My/Ery output and they were thus excluded from further analyses.

Having predicted the best gating strategies to isolate cells with specific *in-vitro* behaviours I characterised both CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells through single cell functional assays. A significant difference in the clonogenic efficiency of these populations was identified, with CD49f$^+$ Subset1 having a higher clonogenic efficiency than CD49f$^+$ Subset2 (Figure 13 B). As hypothesized, CD49f$^+$ Subset1 cells produced significantly more colonies containing Ery cells than CD49f$^+$ Subset2 cells, whereas CD49f$^+$ Subset2 cells made significantly more colonies containing NK cells (Figure 13 C, E). Moreover, CD49f$^+$ Subset1 cells produced a higher proportion of bilineage My/Ery and unilineage Ery colonies than CD49f$^+$ Subset2 cells while on the other hand, CD49f$^+$ Subset2 cells generated a higher proportion of bilineage My/NK and unilineage NK colonies than CD49f$^+$ Subset1 cells (Figure 13 C). When looking at My colonies CD49f$^+$ Subset2 cells produced less My colonies containing both monocytes and granulocytes (MonoGran), but more colonies containing only monocytes (Mon) than CD49f$^+$ Subset1 (Figure 13 D). Division kinetics were also assessed within these prospectively isolated populations and, as predicted, on average CD49f$^+$ Subset1 cells divided significantly later than CD49f$^+$ Subset2 cells (Figure 13 F, G).

In parallel, the same differentiation assays were conducted on Subset1 and Subset2 populations. These yielded similar results to those obtained with CD49f$^+$ HSC (data not shown). Moreover, the colony

output of HSC/MPP subsets was also assessed by CFU- assay and by single cell differentiation assays in condition that allow a better growth of cell of the Ly lineage (B and NK cells) and My lineage through co-culture with a stromal cell line. Results from CFU-assay confirmed that Subset1 cells generate a higher number of Ery-only and mixed (My/Ery) colonies than Subset2 cells while the *in-vitro* differentiation results demonstrated that the Ly differentiation potential of Subset2 cells is not restricted to NK cells but includes production of B cells (data not shown).

In summary, these data demonstrate that single CD49f^{+} HSC or HSC/MPP cells with either My/Ly or My/Ery differentiation outputs *in-vitro* can be prospectively enriched based on distinct levels of CD34 and CLEC9A cell surface expression with CD49f^{+} Subset1 (and Subset1) defined as CD34^{lo} CLEC9A^{hi} and CD49f^{+} Subset2 (and Subset2) as CD34^{hi} CLEC9A^{lo}.

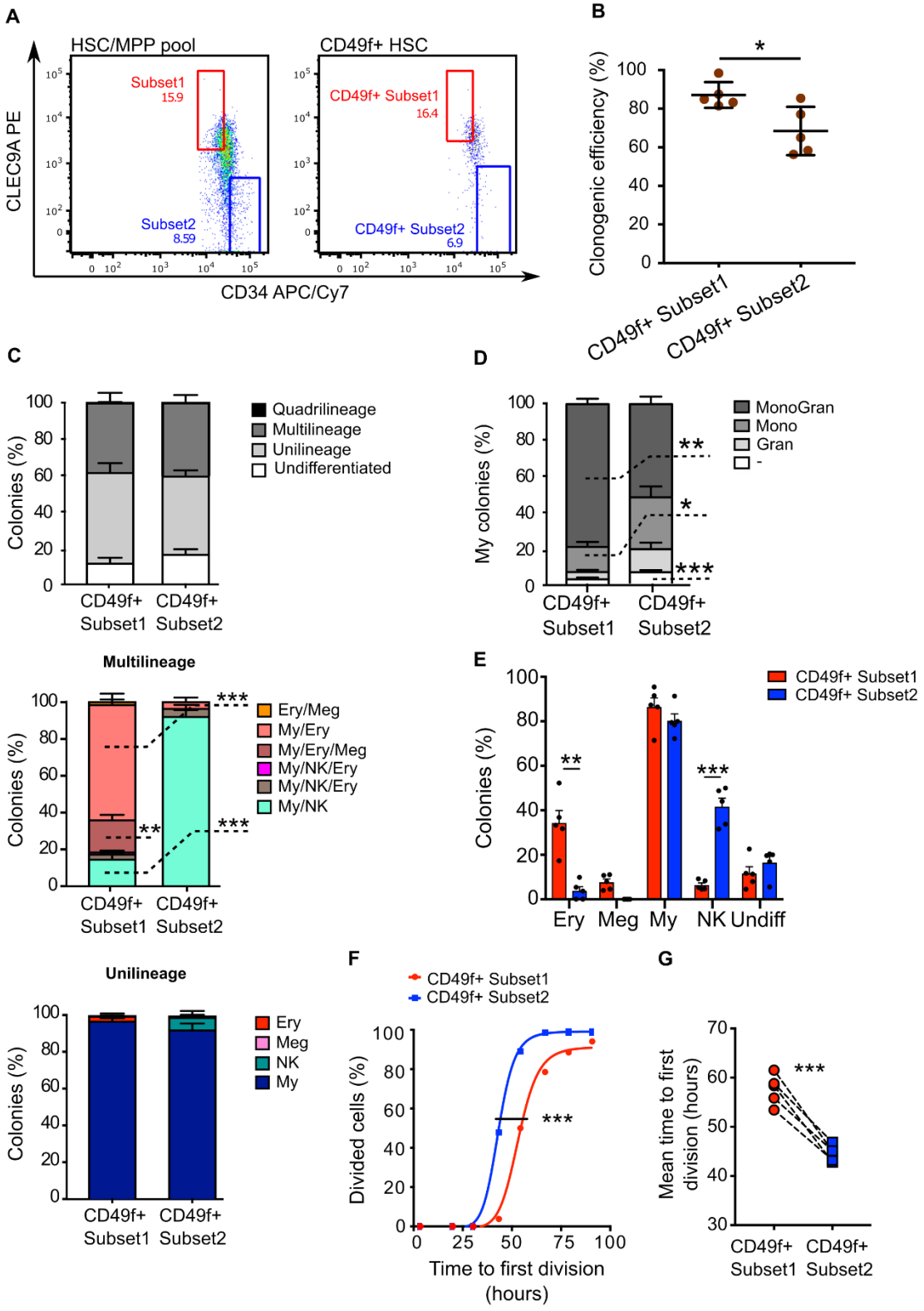


Figure 13: Prospective purification and functional characterisation of CD49f$^+$ Subset1 and CD49f$^+$ Subset2.

A) Representative example of the gating strategies used to isolate the indicated populations as derived by *in-silico* analysis. Percentages of HSC/MPP and CD49f$^+$ HSC within the identified gates are shown. B) *In-vitro* single cell differentiation assay clonogenic efficiency of single CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells (n= 5 independent CB samples, 714 total cells plated for CD49f$^+$ Subset1 and 756 for CD49f$^+$ Subset2, minimum 54 cells per CB). Statistical significance was calculated by two-sided paired t-test (*p<0.05). C, D) Percentage of colonies of the indicated type obtained from *in-vitro* differentiation of single CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells (n =5 independent CB samples,

n = 628 colonies from CD49f$^+$ Subset1 and n = 522 colonies from CD49f$^+$ Subset2). Statistical significance was calculated by two-sided paired t-test (*p<0.05, **p<0.01, ***p<0.001). E) Percentage of colonies generated by CD49f$^+$ Subset1 (red) or CD49f$^+$ Subset2 (blue) single cells containing differentiated cells of the indicated lineages. Mean ± SEM is shown. Statistical significance was calculated by two-sided paired t-test (**p<0.01, ***p<0.001). F) Representative example of first division kinetics for single CD49f$^+$ Subset1 (red) or CD49f$^+$ Subset2 (blue). The graph is a non-linear fit cumulative curve. Statistical significance is calculated by extra sum-of-square F test (***p<0.001). G) Mean time of first division of single CD49f$^+$ Subset1 (red) or CD49f$^+$ Subset2 (blue) (EC50 of non-linear fit of cumulative first division kinetics); n =5 independent experiments with independent CB samples. Statistical significance is calculated by two-sided paired t-test (***p < 0.001).

4.3 Functional and transcriptional characterisation of CD49f$^+$ Subset1 CD49f$^+$ Subset2 HSCs at the single cell level

4.3.1 Transcriptional initiation of lineage priming programmes already occurs within the CD49f$^+$ HSC compartment

After characterising the differentiation output *in-vitro* of the two newly identified CD49f$^+$ HSC subsets I conducted scRNA-seq on CD49f$^+$ Subset1 and CD49f$^+$ Subset2 to investigate their transcriptional status. For each population I index-sorted 96 cells that I then used to perform scRNA-seq. After sequencing, the cells that passed QC were respectively 81% and 78% for CD49f$^+$ Subset1 and CD49f$^+$ Subset2. Data from the sequencing were analysed together with the data obtained from the sequencing of the whole CD49f$^+$ HSC compartment (Figure 12). From a 3D dimensionality reduction analysis (including tSNE, PCA and diffusion map representations) it is possible to see that CD49f$^+$ Subset1 and CD49f$^+$ Subset2 single cells form two separate clusters that occupy opposite poles of the CD49f$^+$ HSC transcriptional space. This supports the idea of the polarised continuum of differentiation derived from the *in-vitro* functional assays and transcriptional analysis of the CD49f$^+$ HSC compartment. (Figure 14 A-D).

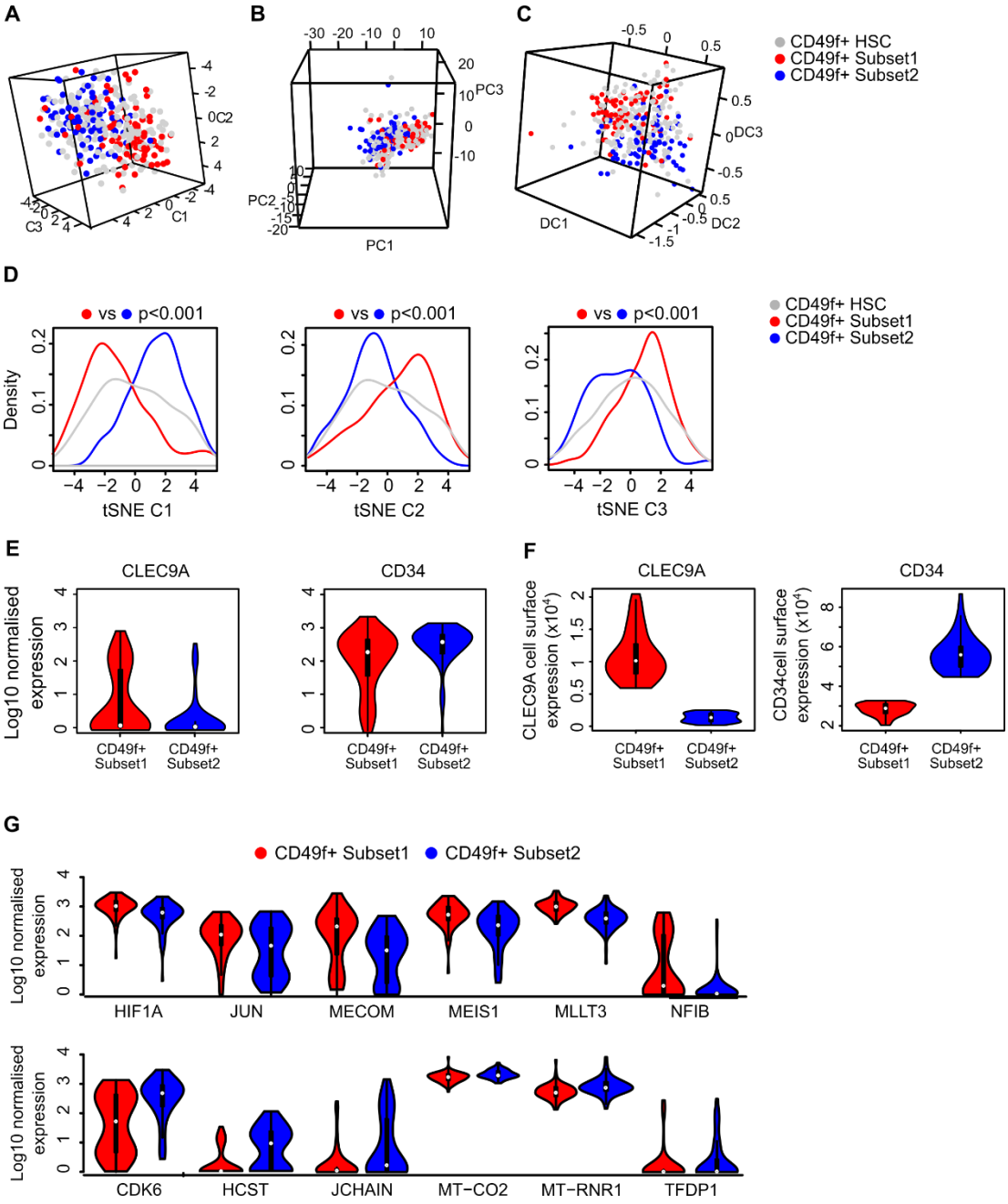


Figure 14: Transcriptional polarisation of the CD49f$^+$ HSC compartment and transcriptional analysis of CD49f$^+$ Subset1 and CD49f$^+$ Subset2.

A-G) ScRNA-seq analysis of CD49f$^+$ HSCs (n = 169), CD49f$^+$ Subset1 (n = 78) and CD49f$^+$ Subset2 (n = 75) cells. A, D) tSNE 3D representation of the transcriptional landscape of CD49f$^+$ HSC (grey), CD49f$^+$ Subset1 (red) and CD49f$^+$ Subset2 (blue) cells. All tSNE analyses were performed on highly variable genes (2420 genes computed as in Brennecke et al. 2013). B, C) PCA 3D (B) and diffusion map (C) representation of single CD49f$^+$ HSCs (grey), CD49f$^+$ Subset1 (red) and CD49f$^+$ Subset2 (blue) cells. PCA and diffusion map analyses were performed on highly variable genes (2420 genes computed as in Brennecke et al. 2013). D) Density plots showing the distribution of single CD49f$^+$ HSCs (grey), CD49f$^+$ Subset1 (red) and CD49f$^+$ Subset2 (blue) cells along the indicated tSNE components. Statistical significance was calculated by two-tailed unpaired t-test comparing CD49f$^+$ Subset1 and CD49f$^+$ Subset2 ($p<0.001$). E) Normalised transcriptional expression (Log_{10}) of CLEC9A (left) and CD34 (right) from 78 CD49f$^+$ Subset1 single cells and 75 CD49f$^+$ Subset2 single cells. F) Cell surface expression measured by index sorting of CLEC9A (left) and CD34 (right) surface markers for the cells analysed with scRNA-seq. G) Normalised expression (Log_{10}) of the indicated differentially expressed genes between CD49f$^+$ Subset1 (red) and CD49f$^+$ Subset2 (blue) cells (FDR < 0.05 by DESeq2).

ICGS algorithm analysis was conducted again this time on the combined dataset including the two subsets and the data from CD49f$^+$ HSC compartment analysis. The algorithm identified three major clusters this time and each with a similar number of cells (Figure 15 A, B). Cluster 1, which also corresponds to Cluster 1 in Figure 12, resulted to be significantly enriched in CD49f$^+$ Subset1 cells whereas Cluster IIa and IIb, which collectively correspond to Cluster II in Figure 12, were enriched in CD49f$^+$ Subset2 cells (Figure 15 A, B). Comparison of the gene expression between the identified clusters showed that the expression of genes which are usually associated with HSC function such as CXCR4 and JUND was higher in Cluster I. Cluster IIb, in contrast, was characterised by the expression of genes related to metabolism and, in particular, mRNA metabolism genes like HNRNPC, PABPC1, DNA replication genes like NAP1L1 as well as HLA molecules. Overall, this indicated that Cluster IIb contains cells with a more active metabolic state at the transcriptional level than Cluster I cells. In comparison to these two clusters, Cluster IIa included cells with heterogeneous expression of the genes identified as specifically expressed in Cluster I and IIb (Figure 15). Correlation of the transcriptomic analysis and CLEC9A and CD34 surface expression resulted to be consistent with data shown in Figure 12 . In fact, single CD49f$^+$ HSCs within Cluster I displayed significantly higher levels of CLEC9A and lower levels of CD34 when compared to single CD49f$^+$ HSCs within Clusters IIa and IIb (Figure 15 D). In summary, when looking at the transcriptional space of CD49f$^+$ HSCs, CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells are in different areas and they differentially express genes linked to HSC function and activation.

Because of the differences in lineage differentiation of these two populations *in-vitro*, lineage-priming programmes may be already established within the CD49f$^+$ HSC compartment. To test this, Gene Set Enrichment Analysis (GSEA) was performed using population-specific gene signatures obtained from highly purified human HSC and progenitor cells and other specific lineage-priming gene modules (Laurenti et al. 2013, 2015; Velten et al. 2017). Through this analysis we identified HSC-related gene sets and gene modules specific for Ery and Meg differentiation enriched in CD49f$^+$ Subset1 cells (Figure 15 C). Consistently, JUN, MECOM, MEIS1 and HIF1A, all genes involved in HSC maintenance, were showing significantly higher expression in CD49f$^+$ Subset1 cells (FDR<0.05; 96 genes Figure 14 G) (Takubo et al. 2010; Kataoka et al. 2011; S.-Y. Lee et al. 2012; Unnisa et al. 2012). Also, MLLT3, the earliest regulator of Ery-Meg differentiation reported to date (Pina et al. 2008), showed higher expression in CD49f$^+$ Subset1. On the other hand, Ly differentiation-related gene sets were found enriched in CD49f$^+$ Subset2 cells (Figure 14 G). In concordance with previous data showing that CD49f$^+$ Subset2 cell are more metabolically active than CD49f$^+$ Subset1, mitochondrial (MT-CO2 and MT-RNR1) and cell cycle genes (CDK6, TFDP1, 35 genes total, Figure 14 G) were found to be expressed at significantly higher levels in this population. CDK6 levels, as previously mentioned in section 1.2.2, directly regulate the length of HSC exit from quiescence (Laurenti et al. 2015). The finding that higher levels of CDK6 are present in CD49f$^+$ Subset2 cells provides a mechanistic principle for the faster kinetics of quiescent exit that are observed in this population (Figure 13 F,G).

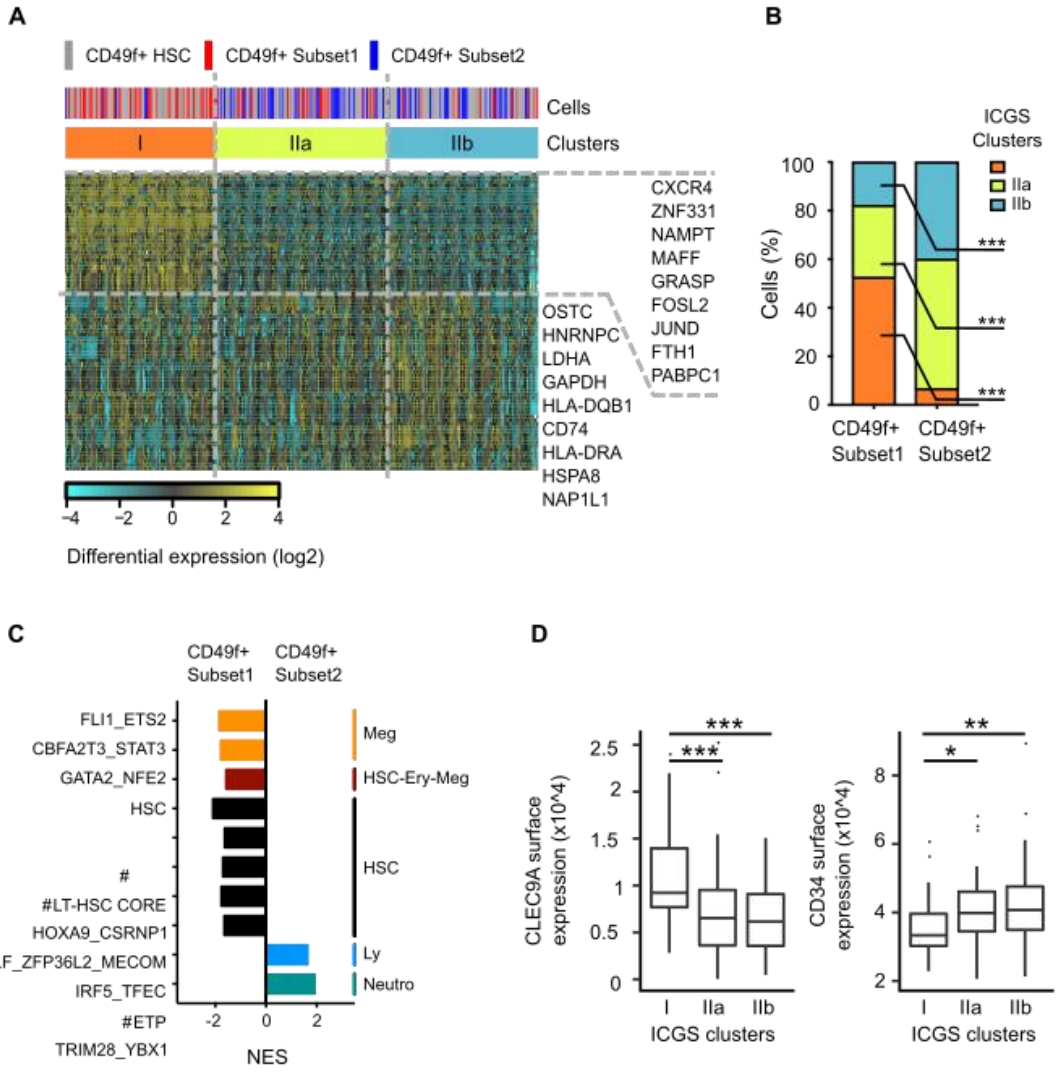


Figure 15: Transcriptional analysis and lineage-priming programme establishment in the CD49f⁺ HSC compartment.

A-C) ScRNA-seq analysis of CD49f⁺ HSCs (n = 169), CD49f⁺ Subset1 (n = 78) and CD49f⁺ Subset2 (n = 75) cells. A) Expression heatmap of guide genes selected by ICGS. Rows represent genes and columns represent single cells. Selected genes are annotated on the side and the cell phenotype is indicated by the uppermost bar of the top panel. The bottom bar of the top panel shows ICGS clusters. B) Percentage of CD49f⁺ Subset1 and CD49f⁺ Subset2 single cells found in each ICGS cluster. Statistical significance was calculated by two-tailed Fisher test (***p < 0.001). C) GSEA analysis results showing significantly enriched (Benjamini-Hochberg adjusted p-value <0.05 by pre-ranked GSEA) population-specific signatures from (Laurenti et al. 2013) and lineage-priming gene modules derived from (Velten et al. 2017) in CD49f⁺ Subset1 and CD49f⁺ Subset2 cells. D) CLEC9A (left panel) and CD34 (right panel) cell surface expression levels in CD49f⁺ HSCs within the indicated ICGS clusters (excluding CD49f⁺ Subset1 and CD49f⁺ Subset2 cells). Median, interquartile and 5-95 percentile are shown. Statistical significance was calculated by one-way ANOVA with Bonferroni correction for multiple comparisons (*p<0.05, **p<0.01, ***p<0.001).

In conclusion, these results showed that the CD49f⁺ HSC transcriptional landscape looks like a continuum of transcriptional states that gradually progresses from a group of HSCs marked by stem cell and early Ery lineage priming programmes (CD49f⁺ Subset1) to another HSC group that shows established Ly-priming and HSC activation programmes (CD49f⁺ Subset2). Moreover, differences in the kinetics of divisions are reflected at the transcriptome level and whether these differences may drive fate choices is still not known.

4.3.2 Differences in the division kinetics of CD49f$^+$ Subset1 and CD49f$^+$ Subset2 are not linked to cell fates decisions

Due to the differences in the division kinetics of single CD49f$^+$ Subset1 and CD49f$^+$ Subset2 and their different expression levels of CDK6 at the transcriptional level, I decided to investigate the link between cell cycle properties and cell fate decisions *in-vitro*. One hypothesis is that, similar to ES cells, the length of the cell cycle may determine differentiation output of HSC subsets. Specifically, here a long time to first division could lead to Ery differentiation whereas a shorter one to Ly differentiation. To test this hypothesis, I took advantage of a specific CDK6 inhibitor, Palbociclib (PD033299 or PD in short) to elongate the time to first division of the faster dividing CD49f$^+$ Subset2 cells and then monitored their differentiation behaviour *in-vitro.* If cell cycle properties are linked to cell fate decisions in HSCs, this assay should reveal an *in-vitro* differentiation behaviour of CD49f$^+$ Subset2 CDK6 inhibited cells similar to the slower dividing CD49f$^+$ Subset1 cells.

Untreated and PD treated (200 nM) CD49f$^+$ Subset2 single cells were sorted and cultured in liquid media as described in Figure 10 A. PD treated cells were washed out after 3 days of culture to reverse CDK6 inhibition and assess their differentiation potential *in-vitro.* CD49f$^+$ Subset1 cells were cultured in parallel and used as a control. The time to first division was monitored and the colony output measured by high-throughput flow cytometry after 3 weeks in culture. After CDK6 inhibition CD49f$^+$ Subset2 cells show a similar time to first division as CD49f$^+$ Subset1 cells confirming that CDK6 plays a central role in the division kinetics of these two subsets (Figure 16 A). Moreover, PD treated CD49f$^+$ Subset2 cells clonogenic efficiency was similar to the untreated counterpart indicating that CDK6 inhibition doesn't affect the *in-vitro* capacity of producing colonies (Figure 16 B). When looking at the type of colonies produced, significant differences were identified in the percentage of colonies containing Ery or NK cells between untreated CD49f$^+$ Subset2 cells and cells treated with PD (Figure 16 C). In addition, as a result of the transient CDK6 inhibition, CD49f$^+$ Subset2 cells produced a significantly higher number of My/Ery and My/Ery/Meg cells compared to untreated cells and a reduced number of My/NK colonies (Figure 16 D). However, this was still vastly different from the distribution of colonies produced by CD49f$^+$ Subset1 cells (Figure 16 D). In summary, even though lengthening the time to first division of CD49f$^+$ Subset2 cells by CDK6 inhibition does slightly impact on their differentiation, it is not sufficient to enforce differentiation properties similar to those of CD49f$^+$ Subset1. As most CD49f$^+$ Subset2 can't produce Ery cells *in-vitro*, and this occurs also independently of their division kinetics, these results suggest that other intrinsic mechanisms may be at the base of their inability to differentiate towards Ery fate.

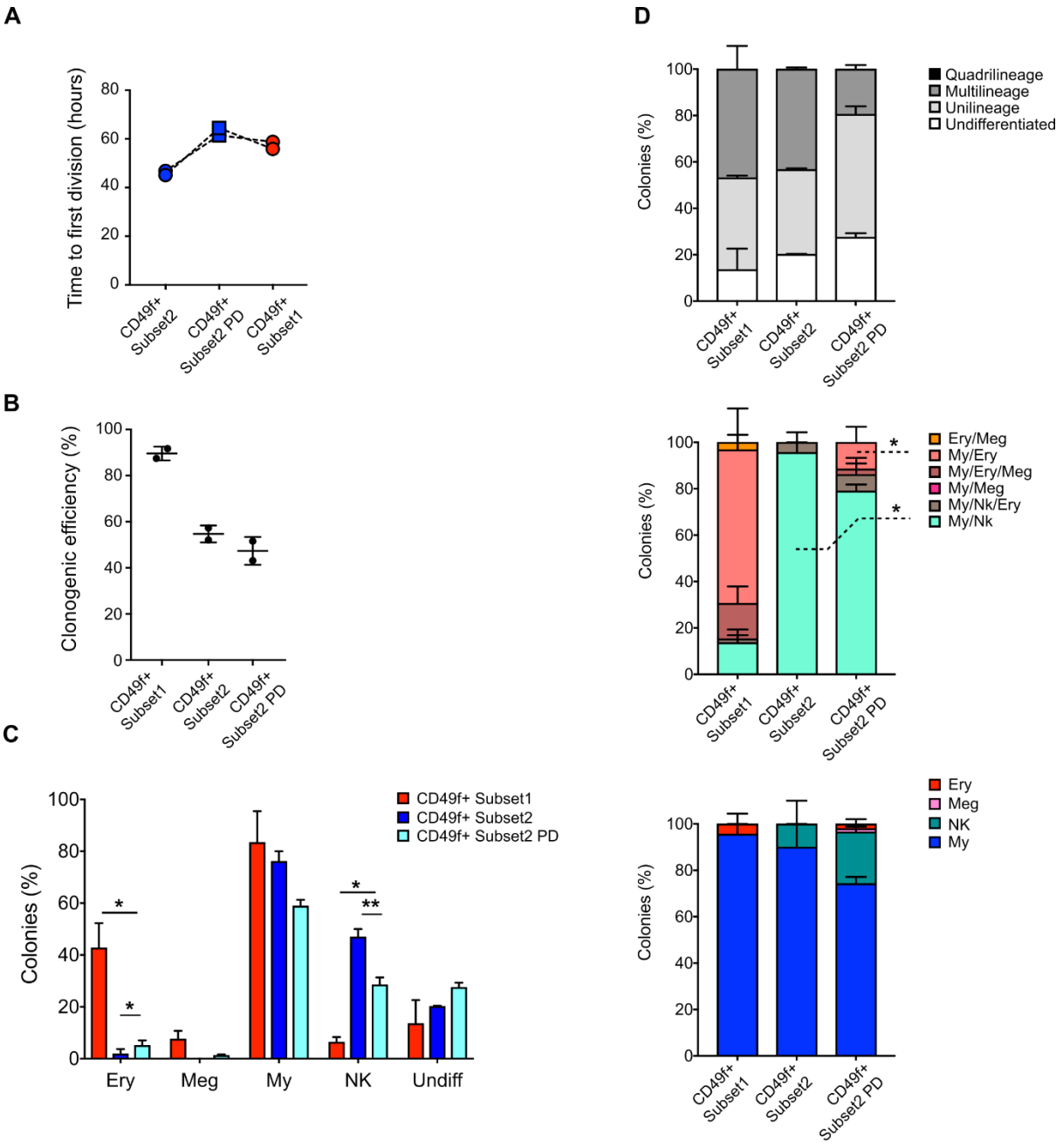


Figure 16: *In-vitro* effects of CDK6 inhibition in CD49f$^+$ Subset2 cells.

A) Mean time of first division of single CD49f$^+$ Subset1, CD49f$^+$ Subset2 and CD49f$^+$ subset2 PD (EC50 of non-linear fit of cumulative first division kinetics); n=2 experiments with independent CB samples. Statistical significance is calculated by two-sided paired t-test (CD49f$^+$ Subset2 UNTR vs PD treated not significant). B) *In-vitro* single cell differentiation assay clonogenic efficiency of single CD49f$^+$ Subset1, CD49f$^+$ Subset2 cells and PD treated CD49f$^+$ Subset2 cells (n= 2 experiments with independent CB samples, 144 total cells plated for CD49f$^+$ Subset1, 336 for CD49f$^+$ Subset2, 192 for CD49f$^+$ Subset2 PD). Statistical significance was calculated by two-sided paired t-test. C) Percentage of colonies generated by CD49f$^+$ Subset1 (red), CD49f$^+$ Subset2 (blue) or CD49f$^+$ Subset2 PD (light blue) single cells containing differentiated cells of the indicated lineages. Mean ± SEM is shown. Statistical significance was calculated by two-sided paired t-test (*p<0.05, **p<0.01). D) Percentage of colonies of the indicated type obtained from *in-vitro* differentiation of single CD49f$^+$ Subset1, CD49f$^+$ Subset2 cells and PD treated CD49f$^+$ Subset2 cells (n =2 experiments with independent CB samples, n = 128 colonies from CD49f$^+$ Subset1, n = 105 colonies from CD49f$^+$ Subset2, n=161 colonies from CD49f$^+$ Subset2 PD). Statistical significance was calculated by two-sided paired t-test (*p<0.05).

4.4 CD49f$^+$ Subset1 cells are hierarchically placed above CD49f$^+$ Subset2 cells

Single cell transcriptomic analyses of CD49f$^+$ Subset1 single cells showed a strong enrichment within this population of HSC specific genes signature. I thus hypothesised that CD49f$^+$ Subset1 cells are positioned above CD49f$^+$ Subset2 cells in the hierarchy and may give rise to CD49f$^+$ Subset2 cells. To identify the hierarchical relationship between these two subsets, both populations were sorted in bulk (Figure 17 A) and cultured in differentiation medium. After 5 days cells derived from the two subsets were then analysed by flow-cytometry and sorted as single cells to perform a secondary differentiation assay. In total, 1152 single cells derived from either CD49f$^+$ Subset1 or CD49f$^+$ Subset2 were index-sorted and placed in liquid culture in the same conditions as in Figure 10 to perform a secondary differentiation assay. Because CD38 and CD90 surface expression is known to vary after culture, these parameters were not included in the analysis. Sorted cells were cultured for 3 weeks and the *in-vitro* lineage output was assessed. Interestingly, when looking at the type of cells originated from CD49f$^+$ subsets at day 5, cells with a phenotype equivalent to that of day 0 CD49f$^+$ Subset2 (CLEC9A^{lo} CD34^{hi}) emerged from cultured CD49f$^+$ Subset1 cells, whereas the parental CD49f$^+$ Subset2 population resulted to be unable to produce cells with the same phenotype as day 0 CD49f$^+$ Subset1(CLEC9A^{hi} CD34^{lo}) (Figure 17 B). Results from the secondary differentiation assay showed that all the populations derived from CD49f$^+$ Subset1 cells (S1_1, S2_1 and Diff1) have a significantly higher clonogenic efficiency than the populations derived from CD49f$^+$ Subset2 cells (S2_2 and Diff2, Figure 17 D) which indicates a more stem-like state of these cells. Of note, levels of CD45RA surface expression were different between Diff1 and Diff2 populations confirming that, despite these populations have similar expression of CD34 and CLEC9A, they are indeed different for the expression of other markers (Figure 17 E). Interestingly, within 5 days in culture CD49f$^+$ Subset1 cells maintained cells with similar differentiation properties as day 0 referred to as S1_1 (Figure 10 E and 17 C, $p>0.1$). This population other than producing the largest colonies (Figure 17 G, $p<0.001$) was also the only population to give rise to multilineage colonies, a property that was suggestive of a more stem-like state similarly to what seen in CD49f$^+$ Subset1 cells (Figure 17 C, top panel). Moreover, CD49f$^+$ Subset1 cells also generated: (i) phenotypically and functionally undistinguishable cells from CD49f$^+$ Subset2, referred to as S2_1, that exclusively produced colonies with My and/or NK cells; (ii) cells with predominant Ery but no Ly output referred to as Diff1 (Figure 17 C-D). On the contrary, at 5 days after culture the progeny derived from CD49f$^+$ Subset2 cells included: (i) cells which produced colonies of small size but qualitatively similar to those of day 0 CD49f$^+$ Subset2 cells although with a higher proportion of NK-cell containing colonies, referred to as S2_2 (Figure 17 C-D); (ii) a smaller population of cells, referred to as Diff2, that originated very few and mostly unilineage colonies (clonogenic efficiency <20%, Figure 17 C-D). My only colonies were further characterised for each of the CD49f$^+$ Subset1 and CD49f$^+$ Subset2 derived population. The results showed that the capacity to produce a MonoGran colonies compared to Mono only or Gran only is higher in S1_1 and it then decreases from to S2_1 and Diff1 while this was not the

case for CD49f$^+$ Subset2 derived populations (Figure 17 F), again suggesting a more stem-like phenotype of CD49f$^+$ Subset1 cells.

In summary, these data are supportive of a model where CD49f$^+$ Subset1 cells are hierarchically placed above CD49f$^+$ Subset2 cells. CD49f$^+$ Subset1 can in fact maintain themselves but also move forward in their differentiation journey and gradually lose their Ery potential generating CD49f$^+$ Subset2 cells.

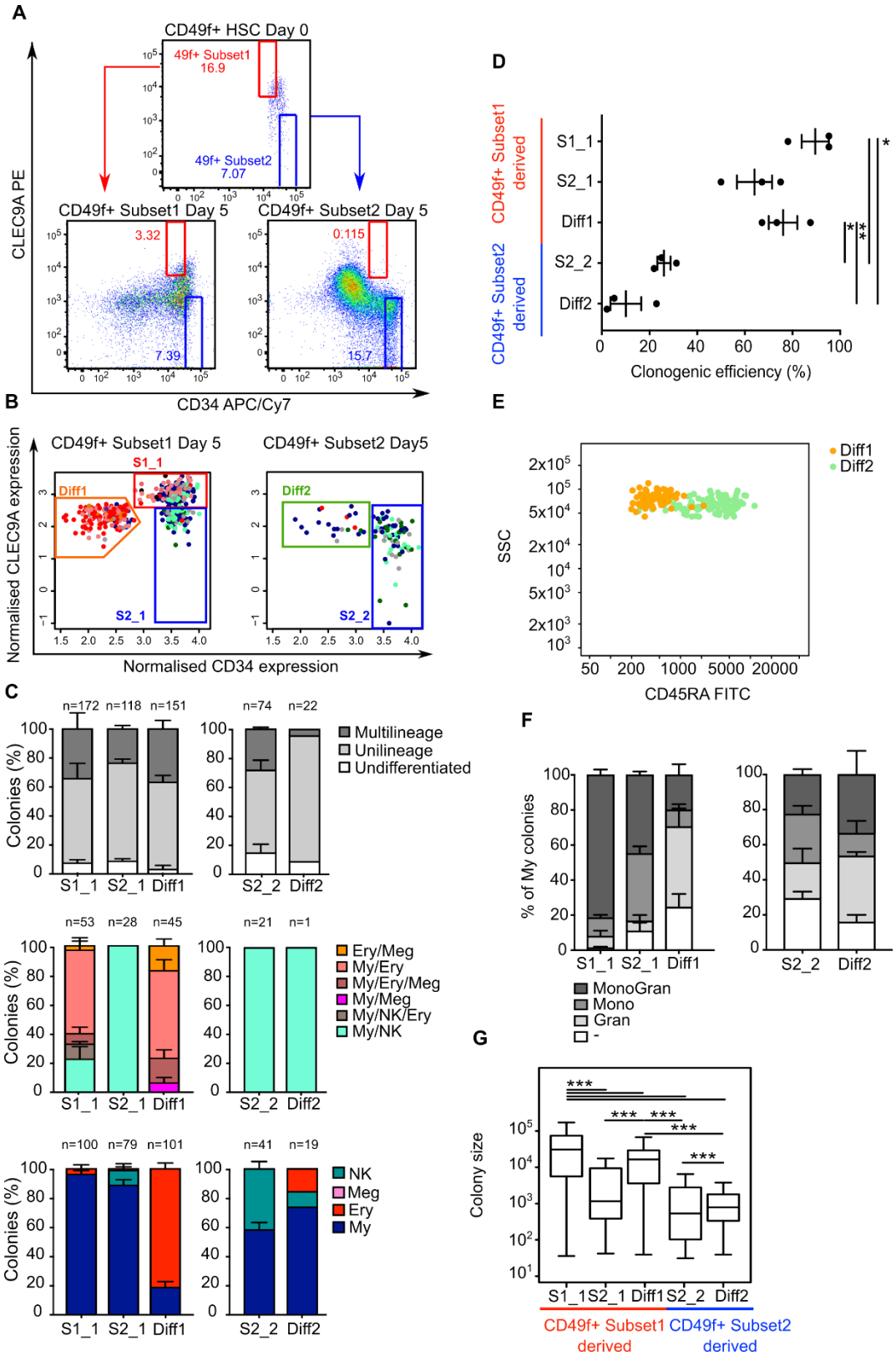


Figure 17: Hierarchical organisation of CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells *in-vitro*.

A) Sorting strategy used to isolate CD49f$^+$ Subset1 and CD49f$^+$ Subset2 populations at day 0 (top) and FACS plot of the derived population after 5 days in culture (bottom). B) Single cells from each of the derived populations at day 5 were sorted. Colours indicate the type of colony produced by each cell after 3 weeks of culture (n = 576 cells were plated from 3 independent CB samples for CD49f$^+$ Subset1 derived populations, n = 576 cells were plated from 3 independent CB samples for CD49f$^+$ Subset2 derived populations) C) Percentage of colonies of the indicated type obtained from *in-vitro* differentiation of single cells from the indicated day 5 populations. (n= 3 experiments from

independent CB samples (except for Diff2 where n =1). Mean ± SEM is shown. The total number of colonies analysed for each population is shown on the top of the bars. D) Clonogenic efficiency (calculated as number of colo
nies/number of seeded cells) of the indicated day 5 populations (n=576 cells plated from 3 independent CB samples for CD49f$^+$ Subset1 derived populations, n=576 cells plated from 3 independent CB samples for CD49f$^+$ Subset2 derived populations). Mean ± SEM is shown. Statistical significance was calculated by one way-ANOVA with Tukey's multiple comparison (*p<0.05, **p<0.01). E) Representative example of CD45RA cell surface intensity for single cells from Diff1 and Diff2 populations. F) Percentage of MonoGran, Mono, Gran and undetermined types myeloid colonies derived from S1_1 (n=159), S2_1 (n=103) and Diff1(n=72); S2_2 (n=62) and Diff2 (n=22) single cells from 3 independent CB samples. Mean ± SEM is shown. G) Size of the different types of colonies generated by single cells from the indicated CD49f$^+$ Subset1 and C49f$^+$ Subset2 derived populations. Size is shown as number of cells per colony. Median, interquartile and 5-95 percentiles are shown. Statistical significance is calculated by Kruskal- Wallis test with multiple comparison (***p<0.001).

4.5 Subset2 cells support durable My/Ly but not Ery engraftment *in-vivo*

As per definition, HSCs can regenerate the entire blood system and thus produce long-term multipotential engraftment after transplantation in the recipient. In the human system, the CD49f$^+$ HSC compartment is the purest compartment isolated to date and can produce robust grafts which last for over 20 weeks after transplantation into NSG mice. While the production of My and Ly lineages is quite efficient in the NSG model, the production of Ery cells after CD49f$^+$ HSC transplantation is poor and not reliable because of the lack of cross-reactivity between the human EPO-receptor and the mouse one. Given their different capacity to generate Ery cells *in-vitro*, to thoroughly assess the repopulating capacity and *in-vivo* lineage potential of CD49f$^+$ Subset1 and CD49f$^+$ Subset2, an improved protocol for *in-vivo* Ery production is needed. A new protocol in which xenotransplanted mice are injected intraperitoneally with human EPO (hEPO) for more than 2 weeks before analysis was established in our lab (experiments performed by Emily Calderbank). The protocol was tested in engrafted mice that were intrafemorally injected with saturating doses of HSCs (CD34$^+$ cells). A 2 weeks treatment with hEPO consistently produced robust Ery grafts in treated mice but not in control mice that were injected with PBS only (Figure 18 B). Ery grafts were measured as percentage of GlyA$^+$ cells in engrafted mice that had total engraftment levels above 0.01%. To ensure that such small grafts in Ery, My and Ly lineage could be reliably detected, specific flow cytometry standards were used (as in methods and Notta et al. 2011). To measure the frequency of repopulating cells within CD49f$^+$ Subset1 and CD49f$^+$ Subset2 while also assessing their lineage potential *in-vivo*, limiting dilution analysis (LDA) experiments 20 weeks post-transplantation were performed and hEPO was administered for 4 weeks before bone marrow analysis (Figure 18 A). The repopulating frequency was calculated using the extreme limiting dilution analysis (ELDA) statistical method (Hu and Smyth 2009). The frequency of repopulation was significantly different between the two subsets ($p=1.06x10^{-10}$): approximately 1:13 cells from CD49f$^+$ Subset1 displayed long-term engraftment capacity at 20 weeks post-transplantation (a similar frequency to that of the unfractionated CD49f$^+$ HSC compartment) (Notta et al. 2011). In contrast, only 1:685 CD49f$^+$ Subset2 cells could produce grafts up to 20 weeks after transplantation (Figure 18 D, E). The lineage composition of the grafts was also analysed and strikingly, Ery cells were detected in the injected bone of 14 out of 19 mice engrafted with CD49f$^+$ Subset1 cells, but in none of the 3 mice engrafted with

CD49f$^+$ Subset2 cells (Figure 18 F). This result supported results obtained *in-vitro* from single cell differentiation assays showing that CD49f$^+$ Subset1 cells mainly produce colonies containing Ery cells (Figure 13C).

It is worth noting that in the NSG model the production of mature lineages is highly skewed towards the B cell lineage as compared to other lineages. However, at 20 weeks post-transplantation 3 mice transplanted with CD49f$^+$ Subset1 displayed unusually low levels of Ly engraftment (marked by arrows in Figure 18 F). These data suggest the presence of a My-bias within CD49f$^+$ Subset1 *in-vivo*.

Altogether, these results show that rare cells with long-term *in-vivo* repopulating capacity are present in both CD49f$^+$ Subset1 and CD49f$^+$ Subset2 populations. However, these long-term repopulating cells are present at different frequencies and display distinct lineage differentiation capacities: CD49f$^+$ Subset1 contains a higher frequency of multipotent (Ly/My/Ery) HSCs whereas CD49f$^+$ Subset2 contains rare long-term repopulating cells, which cannot make erythroid cells but only myelo-lymphoid lineages.

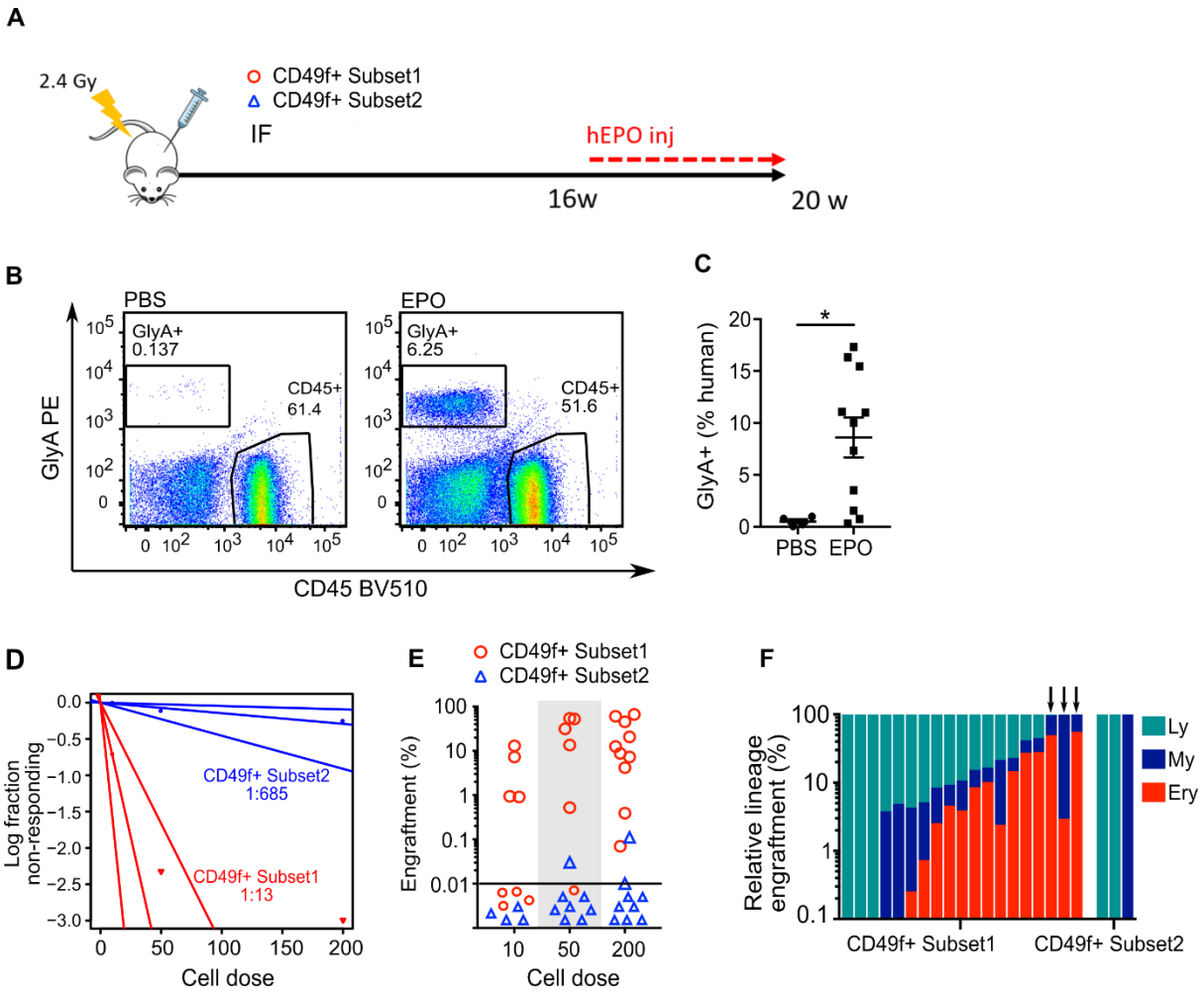


Figure 18: CD49f$^+$ Subset1 and CD49f$^+$ Subset2 show distinct long-term repopulation and differentiation capacities *in-vivo*.

A) Experimental design of the xentransplants of CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells in NSG mice. Cells are injected intrafemorally (IF) 24 h after sublethal irradiation and hEPO is administered intraperitoneally for 4 weeks before BM analysis. B) Representative FACS plot of the engraftment of CB CD34$^+$ cells in the injected femur of control mice injected with PBS only (left) or hEPO (20 units/injection). C) Percentage of total human engrafted Ery (GlyA$^+$) cells in control (PBS, n=4 mice) and treated mice (hEPO, n=11 mice). Mean ± SEM is shown.

Statistical significance is calculated by unpaired Mann–Whitney test (*p < 0.05). D) ELDA estimation of the frequency of long-term repopulating cells within CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells at 20 weeks after transplantation. E) Percentage of human engraftment (% CD45^{++} + % GlyA$^+$) in the injected femur of mice transplanted with CD49f$^+$ Subset1 (red, n= 24 injected mice) or CD49f$^+$ Subset2 cells (blue, n= 21 injected mice) at 20 weeks after transplantation. F) Graft lineage composition analysis. The relative lineage engraftment percentage (for the injected femur) is shown for each engrafted mouse and for each indicated lineage (My, Ly and Ery) 20 weeks after transplantation of CD49f$^+$ Subset1 (n=19 engrafted mice) or CD49f$^+$ Subset2 (n=3 engrafted mice) cells. Bars represent single mice. Arrows indicate mice with unusually low levels of Ly engraftment.

4.6 CD49f$^+$ Subset2 cells are molecularly and functionally distinct from LMPPs

Transplantation assays into NSG mice confirmed that CD49f$^+$ Subset1 cells are multipotent cells with long-term repopulations capacity. On the contrary CD49f$^+$ Subset2 cells with long-term repopulation capacity are rare and unable to produce Ery cells, similarly to My-Ly restricted progenitors such as LMPPs and MLPs (Goardon et al. 2011; Doulatov et al. 2010; Karamitros et al. 2018). However it is important to note that CD49f$^+$ Subset2 cells behaviour *in-vivo* extends well beyond (20 weeks) that of LMPPs and MLPs and, which have only been reported to engraft for up to 2 to 8 weeks after transplantation (Goardon et al. 2011; Doulatov et al. 2010; Karamitros et al. 2018). This suggests that CD49f$^+$ Subset2 cells may contain cells that are placed upstream of LMPPs and MLPs in the haematopoietic hierarchy. To test this hypothesis, I conducted scRNA-seq on 4 different populations including CD49f$^+$ Subset2 cells, Subset2 cells, LMPPs and MLPs. The transcriptional profiles of these populations were compared and visualised using dimensionality reduction tools. By tSNE analysis the transcriptional space was again continuous, with CD49f$^+$ Subset2 and Subset2 cells overlapping but significantly shifted from LMPPs and MLPs (Figure 18 A). Differential expression analysis was conducted. LMPP cells were transcriptionally well distinct from both CD49f$^+$ Subset2 and Subset2 cells (912 genes differentially expressed between LMPPs and CD49f$^+$ Subset2, 437 between LMPP and Subset 2 (Figure 19 B). In contrast, only 50 genes were found to be differentially expressed between LMPPs and MLPs which thus resulted to be transcriptionally very similar. When looking at specific Ly priming specific genes, including BCL11A and SOX4 TFs, which are required for MLPs differentiation into B cells (Laurenti et al. 2013), I found that upregulation of these factors occurred gradually during the progression from CD49f$^+$ Subset2 to Subset2 to LMPP to MLP (Figure 18 C). In parallel, genes that included ID2, involved in the maintenance of HSC self-renewal (Iwasaki and Akashi 2007), or KIT, important for engraftment capacity (Cosgun et al. 2014) were significantly and progressively downregulated from CD49f$^+$ Subset2 to LMPPs/MLPs (Figure 19 C). Moreover, comparison of the expression of cell surface markers like CD34, CLEC9A and CD45RA showed how CD49f$^+$ Subset 2, Subset2 and LMPPs differ in the expression of these markers (Figure 19 E,F,G), again proving that these cells are different from each other not only at the transcriptional level but also phenotypically.

In conclusion, these results show that single cells within the CD49f$^+$ Subset2 and Subset2 compartments represent a developmental intermediate between a multipotent long-term repopulating HSCs and LMPPs. Moreover, loss of Ery potential and Ly lineage priming, although at lower levels than in LMPPs, are already initiated in the CD49f$^+$ HSC compartment within the CD49f$^+$ Subset2, which is characterised

by being erythroid-null, myelo-lymphoid committed cells with infrequent durable reconstitution and faster division kinetics than CD49f^{+} Subset1.

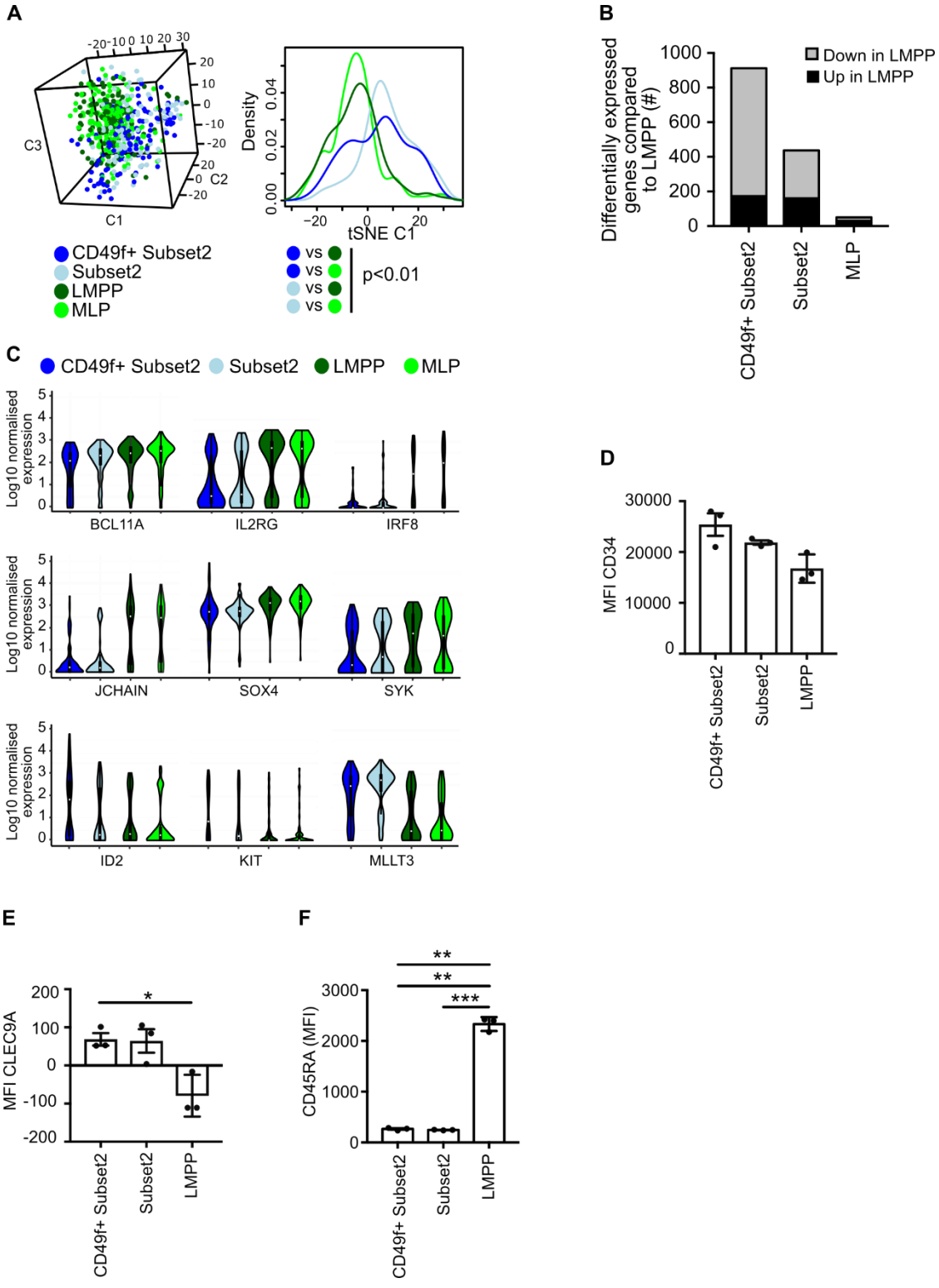


Figure 19: Molecular differences between CD49f^{+} Subset2, Subset2 cells, LMPPs and MLPs.

A) ScRNA-seq was performed on CD49f^{+} Subset2 cells (n=119), Subset2 cells (n=100), LMPPs (n=140) and MLPs (n=134). On the left a 3D tSNE representation of the transcriptional space of the indicated populations, performed on highly variable genes (5831 genes), is shown. The density plot of the distribution of indicated populations along the first tSNE component (tSNE C1) is shown on the right. Statistical significance

was calculated by Kruskal–Wallis test with multiple comparison (*p < 0.01). B) Number of differentially expressed genes between the indicated populations and LMPPs. (FDR <0.05 by DESeq2). C) Normalised expression (Log_{10}) of selected differentially expressed genes both between CD49f$^+$ Subset2 and LMPP and Subset2 and LMPP in single cells from the indicated populations. D-F) Comparison of the cell surface marker expression of CD34 (D), CLEC9A (E) and CD45RA (F) between CD49f$^+$ HSCs, Subset2 cells and LMPPs. The median Fluorescent Intensity (MFI) of each marker for CD49f$^+$ HSCs, Subset2 and LMPPs is shown (n=3 independent CB samples). Mean ± SEM is shown. Statistical significance was calculated by one-way ANOVA with Tukey's multiple comparison (*p<0.05, **p<0.01, ***p<0.001).

Chapter B: Cellular and molecular characterisation of exit from quiescence in human HSCs

4.7 An *in-vitro* model system to study quiescence and activation of human LT-HSCs and ST-HSCs

In order to study the molecular and cellular changes associated with exit from quiescence in human HSCs, and more in detail in the phenotypic LT-HSC (CD49f$^+$ HSCs) and ST-HSC (CD49f$^-$ HSCs) sub-populations, I adopted an *in-vitro* model system based on Elisa Laurenti's publication (Laurenti et al. 2015). In this work the division kinetics of both LT- and ST-HSCs were defined and CDK6 activity was found to be important in the regulation of the duration of quiescence exit. Indeed, CDK6 activity could be blocked using a specific inhibitor (Palbociclib, PD033299 or PD in short) which leads to the lengthening of the time of exit from quiescence in both LT- and ST-HSCs.

This model allows to distinguish between quiescence/exit from quiescence and cell cycle progression during *in-vitro* activation of human HSCs. In this system quiescent LT- and ST-HSCs (q-HSCs) become activated (see definition of activation, section 1.4.1) (a-HSCs) and divide within 72 h in culture in MEM media (see methods). Treating these two populations with the highly specific CDK4/CDK6 inhibitor for 72 h (Palbociclib, PD033299) pushes these cells into a state of pharmacological quiescence (pq-HSCs), either preventing their entrance in the late G_1 phase and thus blocking (LT-HSCs) or partially blocking (ST-HSCs) them in the G_0/early G_1 phase as shown by cell cycle staining in Figure 20. Two different concentration of the CDK6 inhibitor were tested (50 nM and 200 nM): more than 80% LT-HSCs and 50% ST-HSCs were in the G_0/early G_1 phase of the cell cycle after 72 h in culture by Ki67/Hoechst staining (Figure 20). Because the effect of PD was stronger at 200 nM (Figure 20), all the following experiments have been conducted using this concentration.

Thanks to this model I initially studied the molecular changes associated with exit from quiescence and entrance in the cell cycle in human HSCs distinguishing three different conditions: quiescence, activation and pharmacological quiescence.

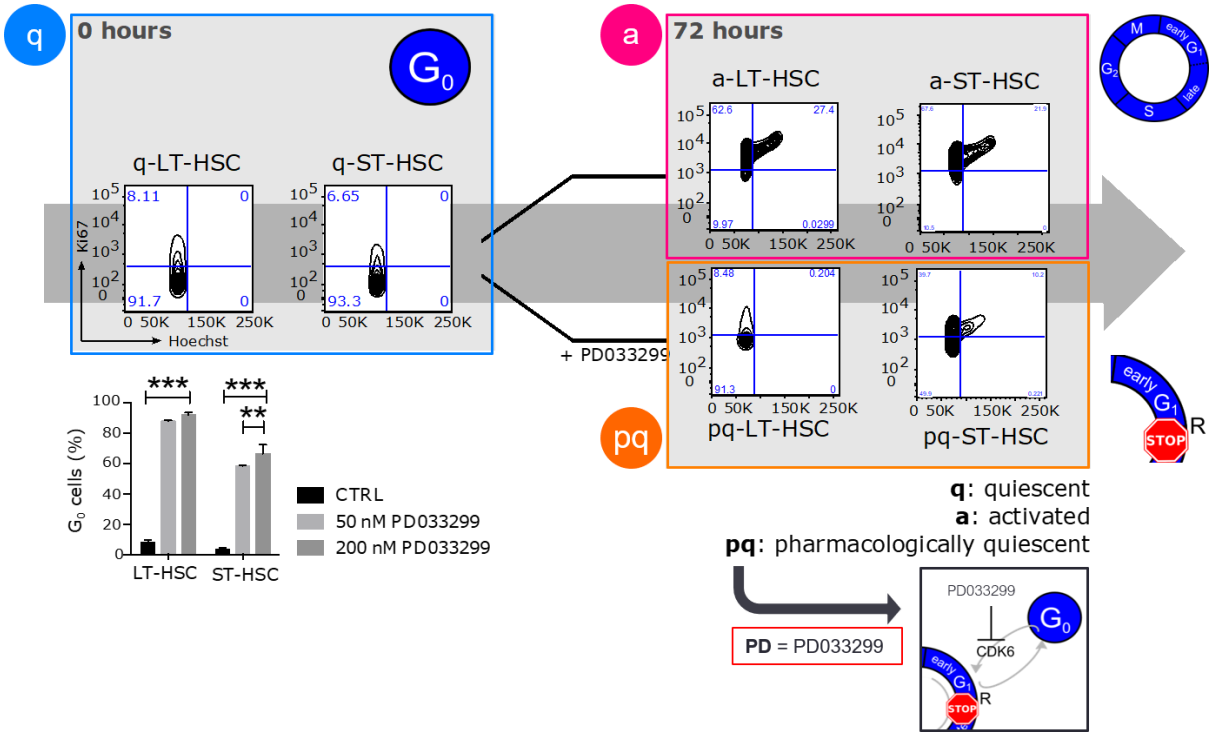


Figure 20: *In-vitro* model system for the study of quiescent and activated HSC.

The top left side of the figure shows representative FACS plots of the cell cycle analysis of freshly isolated (0 hours) LT-HSCs (q-LT-HSCs) and ST-HSCs (q-ST-HSCs). The top right side of the graph shows representative FACS plots of the cell cycle analysis of activated LT-HSCs (a-LT-HSCs) and ST-HSCs (a-ST-HSCs), while the bottom right graph shows the same analysis for pharmacologically quiescent LT-HSCs (pq-LT-HSCs) and ST-HSCs (pq-ST-HSCs). The Ki67/Hoechst staining differentiates the G_0 ($Ki67^-$ $Hoechst^-$), G_1 ($Ki67^+$ $Hoechst^-$), and S-G_2-M ($Ki67^+$ $Hoechst^+$) cell cycle phases. The bar plot shows the percentage of cells in the G_0 phase of the cell cycle in both activated LT- and ST-HSCs (CTRL) and in pharmacologically quiescent LT- and ST-HSCs when treated with two different concentrations of PD033299 (50 nM and 200 nM).

4.8 Characterisation of the division kinetics and proliferation rate of ST and LT-HSCs

With this *in-vitro* model system, I first focused on replicating and extending the division kinetics experiments published in Laurenti et al. 2015 to define HSC division kinetics in cytokine rich media (see methods). A good knowledge of the division kinetics is fundamental to study the regulation of quiescence exit in these cells and best interpret the findings in this model. To do this I cultured HSCs in MEM medium which includes differentiating cytokines and I confirmed that, in these conditions, LT and ST-HSCs complete their first division at significantly different times with an average time of first division of 58 h and 49 h respectively (Figure 21 A-B). These assays also showed that CDK6 inhibition mediated by PD completely stops LT-HSCs but not ST-HSCs from dividing (Figure 21 A). These results are in concordance with the results shown in Laurenti et al. 2015 where ST-HSCs can't be completely stopped from dividing due to their intrinsically higher levels of CDK6 compared to LT-HSCs which, on the contrary, can be almost completely stopped from dividing. Because CDK6 inhibition stops most LT-HSCs from dividing and increases the time of first division in ST-HSCs, it will also affect and thus

reduce the proliferation rate of these populations *in-vitro*. A reduction in cell proliferation over 72 h was indeed identified in both populations as shown in Figure 21 C and D.

To determine the time at which HSCs transition from the exit from quiescence phase (Figure 6) into late G_1 *in-vitro*, the phosphorylation of the RB protein (ser807/811) was assessed at different time points through flow-cytometry (Rubin 2012). This step defines the entrance in late G_1 after which a cell becomes committed to enter the cell cycle and progresses through it independently from external signals. Because the average time of first division in LT and ST-HSCs is much earlier than 72 hours and no cells divide before 24 hours in culture, RB phosphorylation was only assessed at early time-points. Interestingly, as shown in Figure 21 E, RB phosphorylation doesn't happen before 6 h of culture. In fact, at 6 h, less than 10% LT- and ST-HSCs are pRB^+ and most cells phosphorylate RB by 24 h in culture.

Overall, these results provide a solid understanding of the division kinetics and kinetics of exit from quiescence of LT- and ST- HSCs in MEM media and their behavior after inhibition of CDK6, a master regulator of quiescence exit. This will be important to study quiescence exit in these cells and to understand the results that will be presented in the next sections.

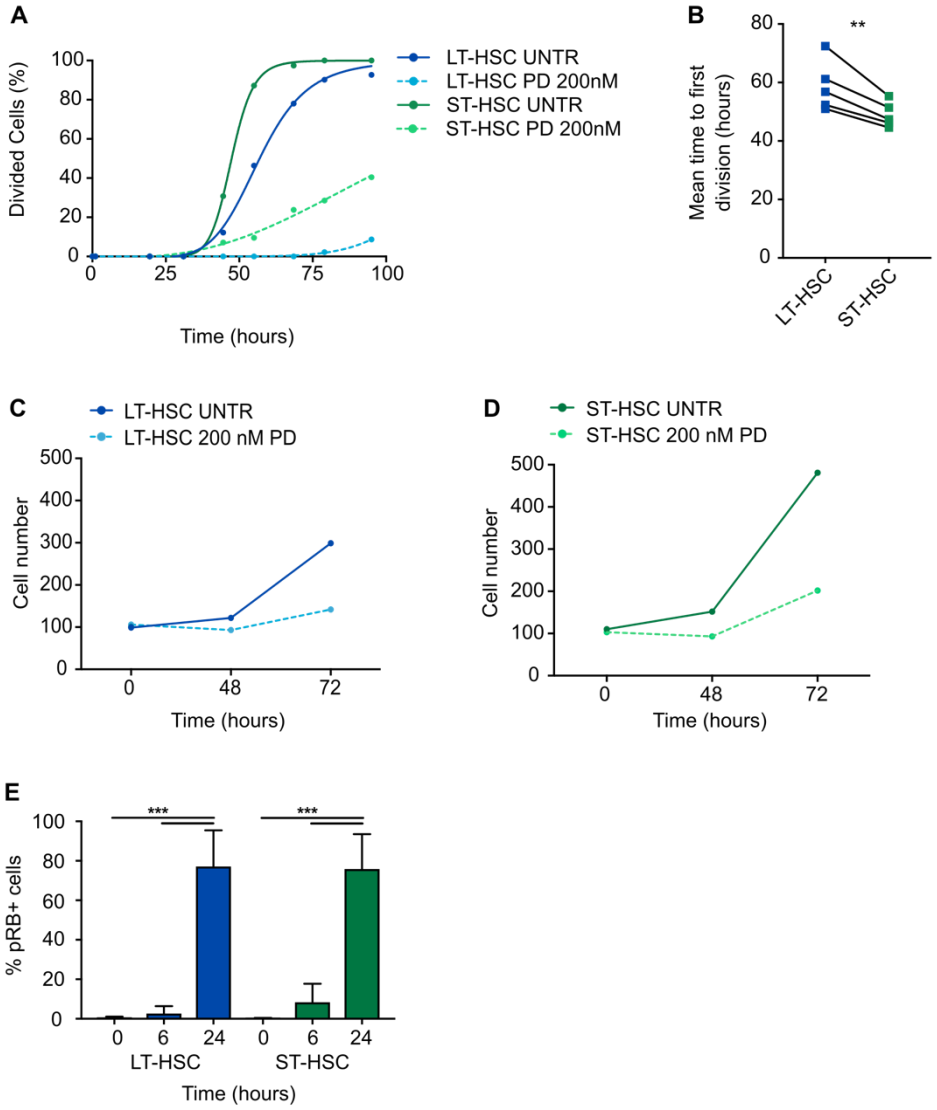


Figure 21: Division kinetics of untreated and CDK6 inhibited LT- and ST-HSCs.

A) Representative example of first division kinetics for single untreated LT- or ST-HSCs and PD treated LT- or ST-HSCs. The graph is a non-linear fit cumulative curve. Statistical significance is calculated by extra sum-of-square F test. B) Mean time of first division of single LT- and ST-HSCs (EC50 of non-linear fit of cumulative first division kinetics); n =5 experiments with independent CB samples. Statistical significance is calculated by two-sided paired t-test (**p<0.01). C-D) Two representative examples of the proliferation rate of 100 LT-HSCs (D) and 100 ST-HSCs (E) over 72 h in culture with and without CDK6 inhibition. The graphs show the number of cells over time. E) Percentages of pRB^{+} cells at different time points are shown. (n=4 experiments with independent CB samples). Statistical significance was calculated by unpaired t-test (***p<0.001). pRB staining assays were performed by Michael Drakopoulos.

4.9 Transcriptomic analysis of bulk RNA-sequencing data for the characterisation of the regulation of quiescence, quiescence exit and activation in LT-HSCs

After having characterised the kinetics of quiescence exit in human HSCs I focused on understanding the molecular regulation of quiescent exit in both LT- and ST-HSCs. To this purpose I analysed bulk RNA-sequencing data obtained from 500 LT- and ST-HSCs at 0h (q), 72 h (a) and 72 h after PD

treatment (pq), accordingly to the model shown in Figure 20. Expression files for each of these conditions were obtained from a collaboration with John Dick's group.

4.9.1 Most transcriptional changes associated with activation occur during quiescence exit in a cell cycle independent manner

Because LT- and ST-HSCs differ in their division kinetics *in-vitro* I first decided to investigate if, transcriptionally, the process of activation occurs in similar ways in these two populations or if there are broad differences in the regulation of activation in these two HSC subsets.

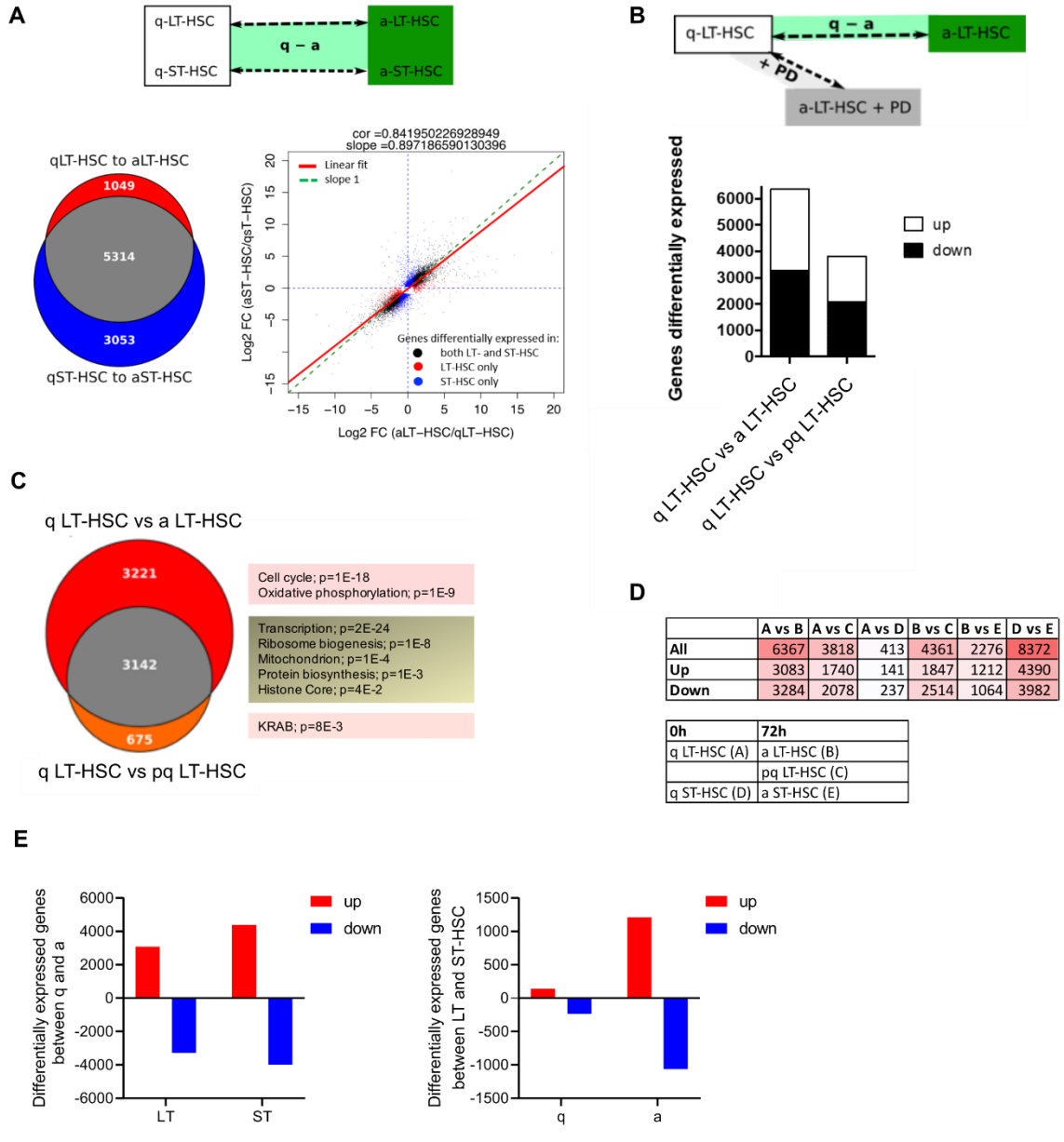
To this purpose, differential expression analysis was conducted by DESeq between q-LT-HSC and a-LT-HSC or q-ST-HSC and a-ST-HSC (Figure 22 A, top). More than 6000 genes were found differentially expressed in both comparisons (6367 and 8372 genes respectively, Figure 22 D and E), with a relatively large overlap between the subsets (Figure 22 A, bottom left). Interestingly, the proportion of up regulated genes and down regulated genes was similar with a ratio of 1:1. Gene expression changes occurring from quiescence to activation were highly correlated between LT- and ST-HSCs (Figure 22 A, bottom right). This overall indicates that the process of activation is very similar between LT- and ST-HSCs at the molecular level. Moreover, as previously reported in Laurenti et al., 2015, only a few hundred genes were found to be differentially expressed between quiescent ST and LT-HSCs (Figure 22 D and E).

Secondly, to compare the molecular programmes involved in exit from quiescence from those associated with cell cycle progression I compared the gene expression data obtained from q-LT- , pq-LT- and a-LT-HSCs (Figure 22 B, top). This comparison was not deemed relevant in ST-HSCs because CDK6 inhibition with PD doesn't completely block these cells from dividing (Figure 21 A). Data obtained from cells treated with PD can only be considered indicative of what happens during exit from quiescence knowing that all the cells are blocked in the early G_1 phase and before PB phosphorylation. Differential expression analysis between quiescent and activated (q vs a), and quiescent and CDK6 inhibited LT-HSCs (q vs pq) revealed that most of the transcriptional changes occur during exit of quiescence and before the RB cell cycle restriction point. Indeed, by DESeq analysis, 3818 genes were found to be differentially expressed in q vs pq. Moreover, in both comparisons, the proportion of up regulated genes was similar to the proportion of the downregulated ones. (Figure 22 B, bottom). These results indicate that gene expression regulation associated with cell culture and *in-vitro* activation already happen during exit from quiescence and it is not dependent on progression of the cell cycle.

To understand which genes and which biological pathways are involved in the regulation of exit from quiescence, a Venn diagram analysis was conducted together with a Gene Ontology (GO) analysis of the differentially expressed genes shared between q-LT- vs a-LT-HSCs and q-LT- vs pq-LT-HSCs. This analysis identified 3142 differentially expressed genes to be regulated during both quiescence exit and activation. GO terms analysis revealed that these genes mainly belong to ribosomal biogenesis and

mitochondria metabolism signatures, other than transcriptional control, protein biogenesis and histone core signatures. As expected, cell cycle signatures were mainly enriched in q vs a and not in q vs pq supporting the idea that expression of cell cycle genes only comes late during the activation process, specifically after exit from quiescence and the G_1 phase restriction point (Figure 22, C).

Altogether, these results show that most transcriptional changes associated with activation occur during quiescence exit and are independent of progression beyond early G_1 and that increased metabolic functions and activation of cell growth pathways are the first responses associated with exit from quiescence.



	A vs B	A vs C	A vs D	B vs C	B vs E	D vs E
All	6367	3818	413	4361	2276	8372
Up	3083	1740	141	1847	1212	4390
Down	3284	2078	237	2514	1064	3982

0h	72h
q LT-HSC (A)	a LT-HSC (B)
	pq LT-HSC (C)
q ST-HSC (D)	a ST-HSC (E)

Figure 22: Bulk RNA-seq analysis of the transcriptional changes associated with activation and exit from quiescence.

A) A scheme illustrating the DESeq analysis comparisons investigated is shown at the top (q vs a). A Venn diagram and a linear fit graph comparing differentially expressed genes between LT- and ST-HSCs quiescent (q) and activated (a) state are shown at the bottom. The number of biological replicates is n=2 for q-LT-HSC, n=3 for a-LT-HSC, n=3 for q-ST-HSC, n=3 for a-ST-HSC. B) A scheme illustrating the DESeq

analysis comparisons investigated is shown at the top (q vs a, q vs pq). The bar graph shows the number of differentially expressed genes and the proportions of upregulated and downregulated genes. C) A Venn diagram comparing differentially expressed genes between quiescent (q) and activated (a) state and quiescent (q) and CDK6 inhibited LT-HSC is shown. On the right the most significant (p-value shown) GO groups identified in the intersection and in the symmetric difference are shown. D-E) The number of differentially expressed genes obtained by DESeq in all the indicated comparisons is shown in the table. Cell colour intensity increases with the number of differentially expressed genes. The total number of differentially expressed genes in shown together with the numbers of upregulated and downregulated genes (D). The number of differentially expressed genes between quiescent and activated LT or ST-HSCs and the number of differentially expressed genes between LT and ST-HSCs in their quiescent or activated state are represented in the graph. Up regulated genes are shown in red and down regulated genes are shown in blue (E).

4.9.2 Patterns of gene expression change during the transition from quiescence to exit from quiescence and activation

To understand how gene expression changes over time from quiescence through exit from quiescence and activation and to identify the biological processes involved in this transition, I used the Short Time-series Expression Miner (STEM) software (Jason Ernst and Bar-Joseph 2006). This software implements the STEM clustering method (J. Ernst, Nau, and Bar-Joseph 2005) which allows to define patterns of gene expression over a short period of time while discriminating for significant (real) and non-significant (random) patterns. Moreover, the biological significance of these patterns can be assessed with Gene Ontology (GO) enrichment analysis for a better interpretation of the results.

In this case, we assumed that the 0 h (q), 72 h PD (pq) and 72 h UNTR (a) could be placed in this temporal order. STEM analysis of this short time-series clustered genes into 16 different model expression profiles of which 8 were statistically significant (coloured profiles, $p<0.05$ before Bonferroni correction). Significant profiles could be grouped into three major profile groups displaying similar trends and identified by the same colour (Figure 23).

Interestingly, association of the profiles with the GO terms revealed that activation of mitochondria metabolism and ribosome biogenesis, and in general pathways related to cell growth are the first response during the process of activation and quiescence exit (profiles 14, 11, 15 and 12, Figure 23 and Table 11). Cellular processes related to cell cycle, chromatin reassembly and cell division are only enriched later in the process of activation (profiles 15 and 12, Figure 23 and Table 11) again suggesting that the transcriptional changes associated with activation are not dependent on cell cycle progression. On the other hand, cellular processes that were down regulated during exit from quiescence included chromatin silencing, positive regulation of transcription and processes related to hematopoiesis and cell differentiation (profiles 4, 0, 2 and 3, Figure 23 and Table 11). Despite some of these were more generic and of more difficult interpretation they suggested the idea that HSC quiescence probably involves the action of many TFs to maintain this apparently “inactive” state of HSCs. Moreover, these results also suggested that the activation process is characterised by a down regulation of genes important in HSC maintenance. Additionally, the fact that genes belonging to the “cell differentiation” GO term are down regulated may suggest that activation of cell differentiation signatures could be important already during

early activation and exit from quiescence rather than later, when HSCs are progressing into the cell cycle.

Despite the oversimplification of the STEM analysis, this allowed to obtain initial important information about the biological changes occurring in HSC from the quiescent state to the active state. Although, a more detailed analysis of the canonical pathways and TFs involved in this regulation is needed and it will be performed by GSEA and presented in the next sections.

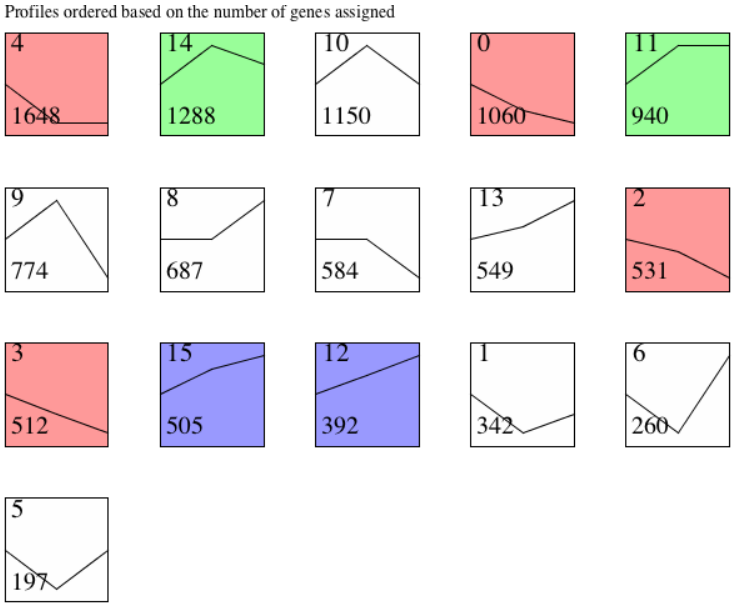


Figure 23: STEM analysis on the 0 h (q), 72 h PD (pq) and 72 h UNTR (a) time series in LT-HSCs.

Each box represents a model expression profile. Profiles are ordered by the number of genes in each profile. Colored profiles have a statistically significant number of genes assigned ($p<0.05$). In each box the profile ID is shown in the top left corner. The lines represent the model profile log of expression change over time. The number of genes assigned to each profile is shown in the bottom left corner. STEM software standard parameters were used for the analysis (see methods).

Profile	GO terms
0	protein modification processes positive regulation of gene expression regulation of cellular metabolic process positive regulation of transcription positive regulation of RNA biosynthetic process
4	regulation of metabolic process positive regulation of gene expression hematopoietic or lymphoid organ development regulation of cell differentiation chromatin silencing
11	mitochondrion ribosome biogenesis rRNA metabolic process oxidation-reduction process tRNA processing intracellular organelle
14	mitochondrion
15	mitochondrion RNA binding ribosome biogenesis mitochondrial envelope ribonucleoprotein complex biogenesis mitochondrial inner membrane oxidation-reduction process

	intracellular organelle
12	mitochondrion protein localization to nucleoplasm mitochondrial membrane cell cycle

Table 11: GO terms significantly enriched in the indicated STEM profiles.

For each of the indicated model profiles, a selection of significant (corrected p-value <0.05) representative GO terms is shown**.**

4.9.3 Quiescence exit is accompanied by remodelling of transcription factors activity

To have a more detailed picture of which pathways are enriched during exit from quiescence (q vs pq) and activation (q vs a) I ran GSEA against the C2 curated gene sets sub-collection of canonical pathways (CP) and the C3 collection of Transcription Factor Targets (TFT) from the MSigDB (Subramanian et al. 2005). Results obtained from the q vs a comparison showed that cell cycle and mitochondria metabolism related gene sets are enriched in a-LT-HSCs. Moreover, other pathways like AURORA B, ATR and PLK1 pathways, involved in post-mitotic surveillance and G_2/M transition, were also enriched in a-LT-HSCs. On the other hand, inflammation related gene sets including interferon (IFN), NFAT, AP1 and CD40 pathways resulted to be the prevalent signatures enriched in q-LT-HSCs. GATA3 and RHOA regulation pathways were also among other interesting pathways enriched in q-LT-HSCs (Figure 24 A).

A heatmap of the top 50 genes expressed within each phenotype (q and a) is shown in Figure 24. Among the highly expressed genes in the quiescent state there are NR4A1 and NR4A3 genes that were recently confirmed to have an important role in restricting HSC proliferation (Freire and Conneely 2018). Other genes included FOSB and JUND, that together form the TF complex AP1 and MAFF and EGR3, two genes with opposing effects on cell cycle, that have been reported to be highly expressed in HSCs where they cooperate to regulate cell cycle (Li et al. 2017). Cell cycle related genes including E2F, CDK1 and CCNA2 and genes of the cytochrome P450 family such as CYP1A1 and CYP2S1 were instead included in the top 50 features of the activated phenotype (Figure 24 B).

According to the results already presented in Figure 22, most of the transcriptional changes associated with activation already occur during exit from quiescence. Here, I have investigated the pathways enriched in q vs pq to understand which ones are activated first during quiescence exit. Results from GSEA analysis showed enrichment of several metabolic pathways (Figure 24 C) including nucleotide metabolism and mitochondria metabolism but, as expected, no cell cycle related signatures. However in pq, I found genes involved in metabolism like genes from the Cytocrome P450 family, and the growth promoting TF MYC which is known to regulate the balance between self-renewal and differentiation in HSCs (A. Wilson et al. 2004; Laurenti et al. 2008). This analysis supports the results from STEM analysis, overall indicating that metabolic activation of HSCs is the first response associated with quiescence exit, followed by cell cycle activation.

The STEM analysis reported in section 4.9.2 supports a model in which a series of TFs have an active role of TFs in the maintenance of quiescence. To assess more in details the activity of the TFs involved in the regulation of quiescence and quiescence exit, I ran GSEA against the transcription factor targets (TFT) signature database. After combining all the results obtained from the comparison of q vs a and q vs pq I generated the graph shown in Figure 25. Three TF activity patterns can be identified:

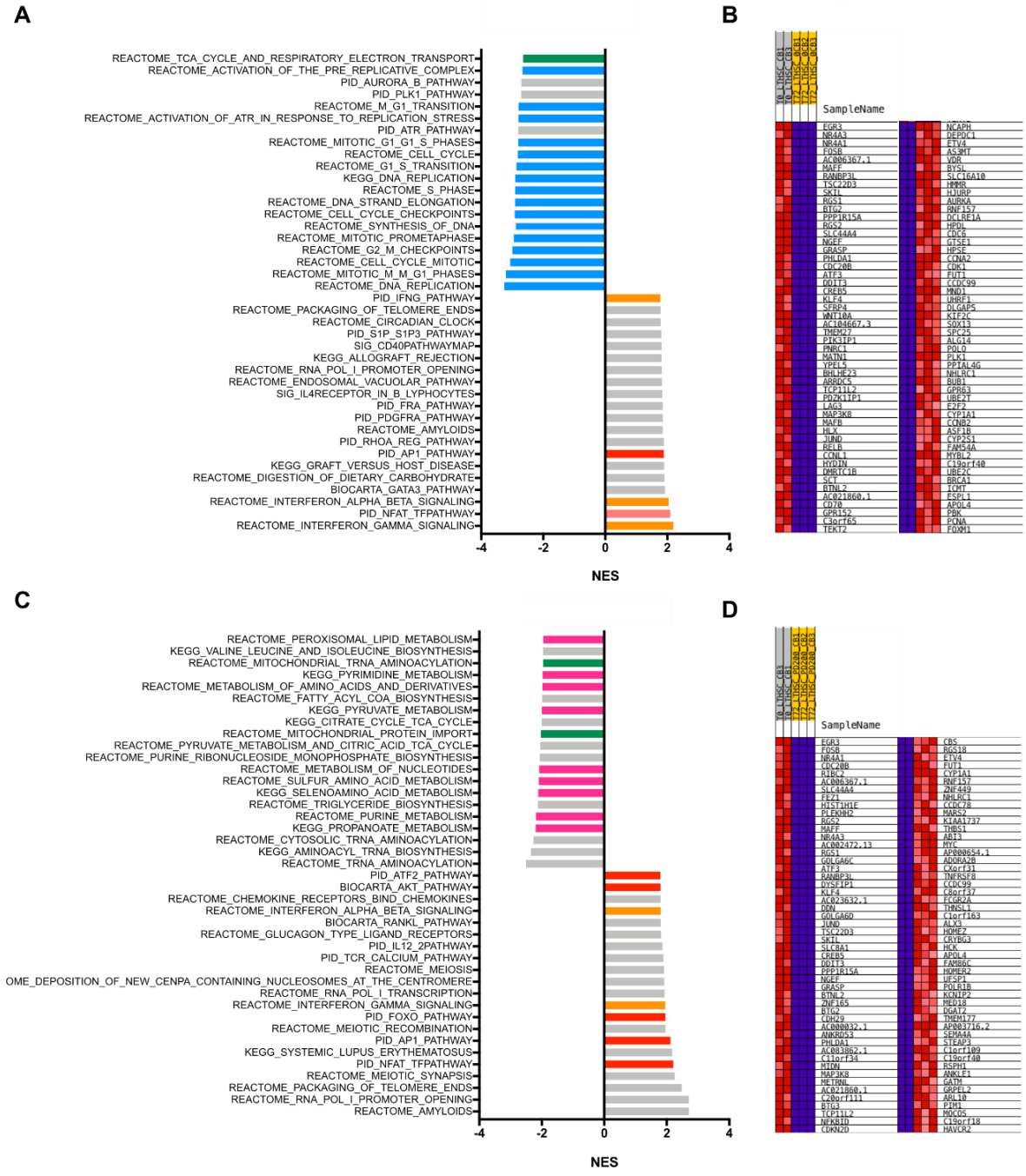
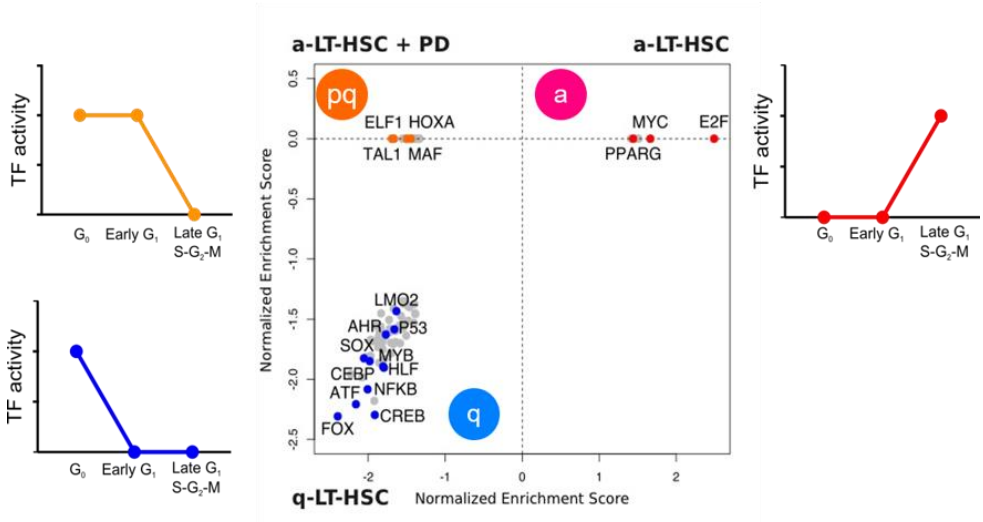


Figure 24: GSEA analysis for common pathways (CP) signatures on q vs a and q vs pq comparisons in LT-HSCs. A, C) GSEA analysis results showing significantly enriched (FDR<25%) common pathways (CP) signatures in the q vs a LT-HSC comparison (A) and q vs pq LT-HSC comparison. Similar colours indicate gene sets of similar biological processes. Gene sets indicated in grey and they haven't been assigned to a specific macro group. B, D) Heat Map of the top 50 features for each phenotype in q vs a comparison (B) and q vs pq (D) LT-HSC comparison. Columns represent different samples.

I) TFs whose activity is rapidly downregulated already during exit from quiescence, hence postulated to play an active role in quiescence maintenance (Figure 25, bottom left, blue); II) TFs whose activity is maintained during quiescence exit but is not required for transition to late G_1-S-G_2-M (Figure 25, top left, orange); III) TFs whose activity is upregulated only after exit from quiescence from late G_1 onwards (Figure 25, top right, red).

Interestingly, these results show that quiescence maintenance is possible thanks to the activity of many TFs while activation appears more as a default cellular state driven by the activity of TFs such as MYC, E2F and PPARG. This again highlights how quiescence is not, as it was previously thought, an inactive state of HSCs but is rather a phase that needs to be carefully regulated in order to be maintained. Among the TF active in quiescence, I identified LMO2, AHR, P53, MYB and HLF, which have previously been reported to be involved in quiescence maintenance in HSCs (Cleveland et al. 2013; Bennett et al. 2015; Liu et al. 2009; Baker et al. 2014; Komorowska et al. 2017). TF targets for ATF and CREB, important TFs for the maintenance of normal haematopoiesis (Cheng et al. 2008) were also enriched in the quiescent state. While the activity of most TF involved in HSC maintenance and quiescence is rapidly downregulated during exit from quiescence, for some of these factors, including HOXA, TAL1 and MAF, this only occurs after entrance in late G_1 (Figure 25).



Taken together, these results show that TF activity is predominantly changed during G_0 exit rather than later in the activation process, demonstrating how, to maintain the quiescence state in HSCs, a high regulated molecular network is required.

Figure 25: GSEA analysis for transcription factor target (TFT) signatures on q vs a and q vs pq comparisons in LT-HSCs.

The graph shows the results of a combined comparison of the GSEA analysis results obtained from enrichment analysis of transcription factor targets signatures (TFT) in the q vs and q vs pq LT-HSC comparison. TF activity trends are shown for each cell state. The most significant TFT signature are highlighted.

4.10 A single cell time-course analysis of the transcriptional landscape of LT-HSCs from quiescence to activation

The transcriptional analysis just presented in bulk LT-HSCs showed that most transcriptional changes related to activation already occur during exit from quiescence. However, in this analysis, the events occurring during quiescence exit were inferred from cells pharmacologically stopped at the end of early G_1 for a significant period in culture. As a result, it is difficult to establish if the changes observed are related to the time spent in culture or the block in the G_0 exit state (G_0/early G_1 phase). Analysis of earlier time points in presence or absence of CDK6 inhibition is thus required. Moreover, transcriptomic analysis of bulk samples only gives an averaged information of the gene expression which may not be representative of what occurs in each of the cells in the population analysed. For these reasons, I performed a time-course scRNA-sequencing on LT-HSCs to investigate the transcriptional changes associated with exit from quiescence and initiation of the activation process at the single cell level. I thus cultured LT-HSCs in low cytokine media (see methods) for 6, 24 and 72 h and 24 h and 72 h with PD and index-sorted single cells at each of these time points to perform scRNA-seq. The results shown here are obtained from the analysis of a total number of 835 single LT- HSCs (539 of which passed QC; 64% efficiency) derived from two independent experiments which, for the purpose of this analysis were combined bioinformatically (with appropriate batch correction – see methods).

To identify the transcriptional changes associated with quiescence exit and activation in correlation to the time spent in culture, a first analysis was conducted on the time-course data obtained at 6, 24 and 72 hours without CDK6 inhibition. Dimensionality reduction on HVGs through 3D UMAP analysis showed that changes at the transcriptional level mostly occur gradually over time (Figure 26 A). However, whereas a trajectory of transcriptional changes was evident from 0 h to 24 h to 72 h, the 6 h time point emerged outside of this trajectory, potentially indicating a specific different state of LT-HSCs at this time point. Because previously presented functional data showed that RB phosphorylation occurs between 6 and 24 h in culture in these conditions (Figure 21 E), I expected the 6 h timepoint to be representative of the exit from quiescence phase. Cell cycle phase assignment for all single cells analysed confirmed that almost 100% and more than 50% of the cells are in the G_0/G_1 phase at 6 h and 24 h respectively as expected, while most cells at 72 h were assigned to any other phase of the cell cycle (Figure 26 C and D). Differential expression analysis with DESeq2 showed that a large number of transcriptional changes already occur within 6 h, and thus during exit from quiescence, with most genes being up regulated rather than down regulated. Interestingly, 38% of the genes differentially expressed at 72 h are already differentially expressed after 6 h in culture (Figure 26 B) and the number of differentially expressed genes increases with the time spent in culture and cell cycle progression (Figure 26 B and D). Overall this indicates that exit from quiescence is associated with gradual changes in expression of genes associated with HSC activation. But many genes are only differentially expressed at 6 h and not later on (2175 genes differentially expressed from 0 to 6 h but at no other time point),

suggesting that this time-point during quiescence exit may represent a specific state of HSCs required before further progressing into the cell cycle.

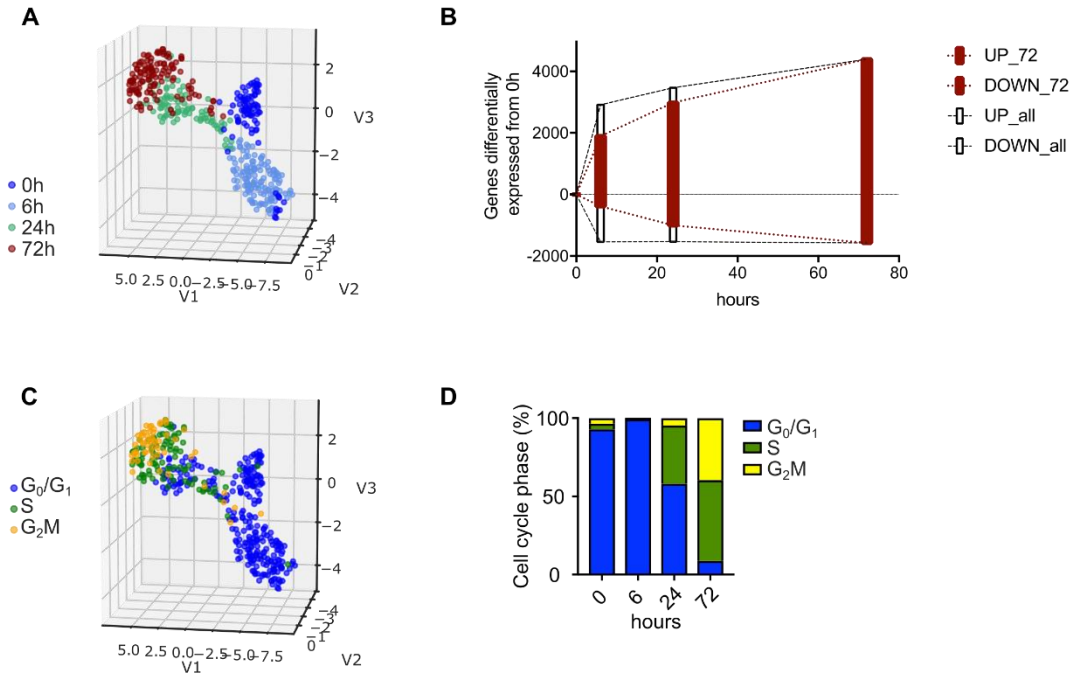


Figure 26: Time course scRNA-seq on single LT-HSCs.

A-D) ScRNA-seq analysis of LT-HSCs cultured for 0, 6, 24 and 72 h (0 h n=85; 6 h n=134; 24 h n=86; 72 h n=124 cells). A) 3D UMAP representation of indicated populations performed on highly variable genes (16276 HVGs). Different colours indicate cells analysed at different time points. B) Number of genes differentially expressed from 0 h to the indicated timepoints are shown (positive axis: upregulated from 0 h, negative axis: down regulated from 0 h). Red bars indicate differentially expressed genes also differentially expressed from 0 h to 72 h. C) 3D UMAP representation as in (A) indicating the cell cycle phases assigned to each single cell. D) The bars indicate the percentage of cells in each of the indicated cell cycle phases for each timepoint analysed.

4.10.1 Gene expression patterns clustering analysis of LT-HSCs time-course scRNA-seq data

To assess how gene expression changes over time and what are the main biological processes and genes involved, a gene expression pattern clustering analysis was performed on the 0, 6 ,24 and 72 h dataset. A total number of 10010 differentially expressed genes between any two time points was identified and these clustered into 11 different patterns (only patterns with a minimum number of 15 genes were considered) (Figure 27). Figure 27 shows all patterns lined by similarity. Clusters type can be mainly subclassified in patterns of up regulation (clusters 1, 2) and patterns of down regulation (clusters 7, 8) over time or a mix of both (clusters 3, 4, 5, 6, 9, 10, 11), where some genes are first either up or down regulated at 6 h and then vice-versa.

ClueGO analysis was performed for each cluster (cluster 6 was excluded due to the low number of genes in the cluster) or for combined clusters, if very similar between each other's, to visualize biological terms (GO terms) genes in a functionally grouped network. Priority of analysis was given to those

patterns of gene expression where genes are down or up regulated right after 0 or 6 h and where this regulation would persist over time (clusters 1, 2, 7, 8). This, because these patterns are more representative of important changes related to exit from quiescence and activation.

Interestingly, some of the ClueGO analysis results for these clusters were similar to what identified with STEM analysis or GSEA analysis of bulk RNA-seq data. In particular, ribosome, mitochondria and oxidative phosphorylation related processes were found to be gradually up regulated during exit from quiescence and activation in cluster 2 (Figure 27). Ribosome and HSF1 related processes were also up regulated during exit from quiescence in cluster 1. These included several ribosomal protein coding genes of the RPL and RPS (S and L ribosomal proteins) families (RPS13, RPS16, RPS23, RPL27, RPL30, RPL35 among others), and heat shock proteins of the HSPA family (HSPA1A, HSPA1B, HSPA2, HSPA8). Interestingly, activation of the p38MAPK cascade was identified at later time points during HSC activation. This could indeed be linked to stress response due to the presence of cytokines in the culturing media but also, it is in accordance with other studies showing that *ex-vivo* expansion of human CB $CD133^{+}$ cells activates the p38MAPK cascade. Oxidative stress has also been liked with activation of this signalling pathway and overall, this is in accordance with the data shown in Figure 27 where activation of p38 MAPK correlates with the time spent in culture and oxidative stress that occurs during metabolic activation (J. Zou et al. 2012)

On the contrary, when looking at which processes were down regulated through exit from quiescence and activation, I identified MHC class II gene sets (including HLA-DRA, HLA-DRB1 and HLA-DRB5) (cluster 8), ribosome processes (cluster 7), although including different RPL and RPS genes from the ones identified previously in cluster 1 and 2. Hematopoiesis and IFN signaling related pathways, which I previously found to be enriched in the 0 h time point in the GSEA analysis results presented earlier (Figure 27, Figure 24 for GSEA analysis), were here enriched in clusters 5 and 7. Among the genes rapidly down regulated during quiescence exit and associated with haematopoiesis FOS, JUN, JUNB, CD74, GATA3, FOXP, HLX and HIF1A were identified. Some of these were previously mentioned as they were also found up regulated at 0 h in bulk RNA-seq data. Other genes that were gradually down regulated included many HLA molecules such as HLA-A, -B, -C and -F and EGR1 and several IRF molecules including IRF4 and IRF7.

Taken together, these results show that genes are regulated in a different number of ways during exit from quiescence to progression in the cell cycle, although an important part of these genes is already up or down regulated during exit from quiescence. Once again, exit from quiescence is associated with rapid downregulation of genes important for HSC and quiescence maintenance and upregulation of metabolism and cell growth related genes. Overall, transcriptomic analysis of single cells provides very similar results to data obtained in bulk, confirming that these responses are specific for all LT-HSCs and no differences can be seen at the single cell level. However, from these data, the 6 h timepoint resulted to be a particular state of HSCs where the cells probably prepare to exit quiescence and thus activates immune response pathways in response to external cytokine stimuli that are present in the culturing

media. A general metabolic and cell growth response is instead what characterises quiescence exit and progression in the cell cycle, even before activation of cell cycle related genes.

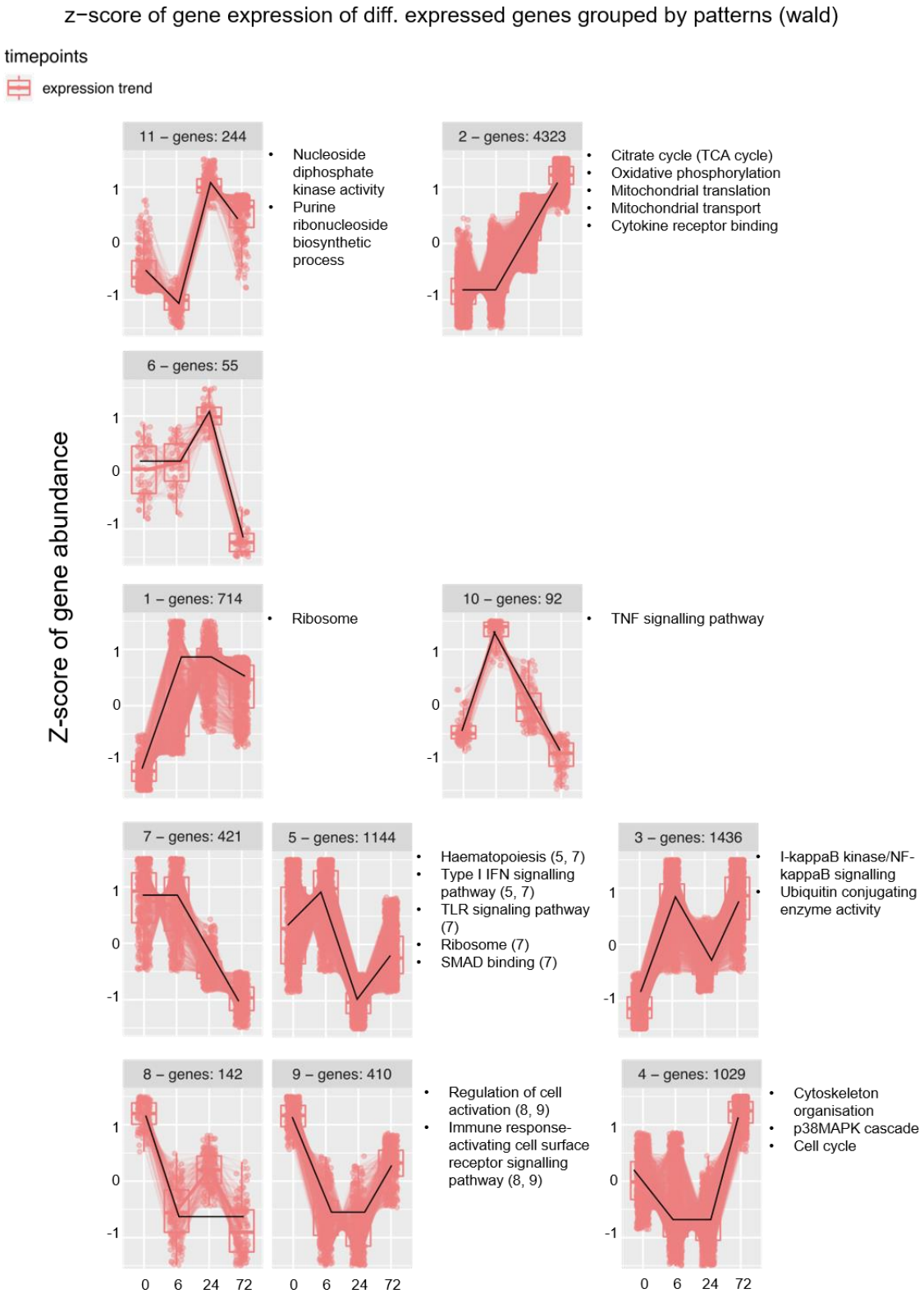


Figure 27: Patterns of gene expression identified during time course of LT-HSC activation.

Identified clustering patterns are shown together with pattern number and number of genes in the cluster. For each time-point the gene expression trend is shown by the black lines. Red lines indicate expression of single genes within the cluster. Selected significantly enriched (Bonferroni adjusted p <0.1) GO terms identified with ClueGO are shown for each expression pattern.

4.10.2 LT-HSC exit from quiescence and activation occurs independently from cell cycle progression

A second analysis was performed combining the previous dataset with data from untreated (UNTR) single cells cultured for 24 h and 72 h in presence of PD (24 h PD and 72 h PD). The aim was to investigate if the transcriptional changes occurring over 72 h in culture are dependent on cell cycle progression and to assess potential transcriptional differences between LT-HSCs stopped at the end of early G_1 and those progressing into the cell cycle at different time points. Once again, dimensionality reduction analysis identified the 6 h as a clearly separate group that doesn't follow the transcriptional trajectory identified for the other time points (Figure 28 A). Moreover, PD treated LT-HSCs from both 24 and 72 h were similarly positioned in the transcriptional space and most of the cells within these groups resulted, as expected, to be in the G_0/G_1 phase of the cell cycle (Figure 28 B) (90% of 24 h PD cells and 74% of 72 h PD cells, data not shown)

Pseudotime analysis was conducted to build temporal trajectories based on the transcriptional states of the cells identified. As expected, 0 h, 6 h, 24 h UNTR and 72 h UNTR single cells were placed in succession along the pseudotime trajectory. Similar to what observed with the dimensionality reduction techniques, 24 h PD and 72 h PD substantially overlapped with 24 h UNTR cells along this pseudotime, indicating a high degree of similarity in the transcriptional changes of these cells (Figure 28 C). Importantly, when pseudotime analysis was repeated removing cell cycle genes (see methods), it yielded almost identical results, suggesting that most transcriptional changes are cell cycle independent.

I then compared genes differentially expressed between 24 h PD and 24 h UNTR cells as well as 72 h and 72 h PD. Compared to their respective UNTR cells, 238 genes were significantly downregulated in 24 h PD and 659 in 72 h PD ($p.adj<0.05$). Pathway analysis identified, as expected, cell cycle genes to be enriched in both gene sets (predominantly related to G_1/S transition at 24 h, and to all phases of the cell cycle at 72 h, data not shown). 286 genes were significantly upregulated in 24 h PD compared to 24 h UNTR and 105 genes in 72 h PD vs 72 h UNTR ($p.adj<0.05$). Of note, ribosomal proteins were highly enriched in the latter.

Altogether, the single cell analysis of LT-HSC treated with PD supports the results obtained from bulk RNA-seq. I conclude that most of the transcriptional changes occurring during *in-vitro* culture of LT-HSC are independent of cell cycle progression and are rather due to exposure to activation signals.

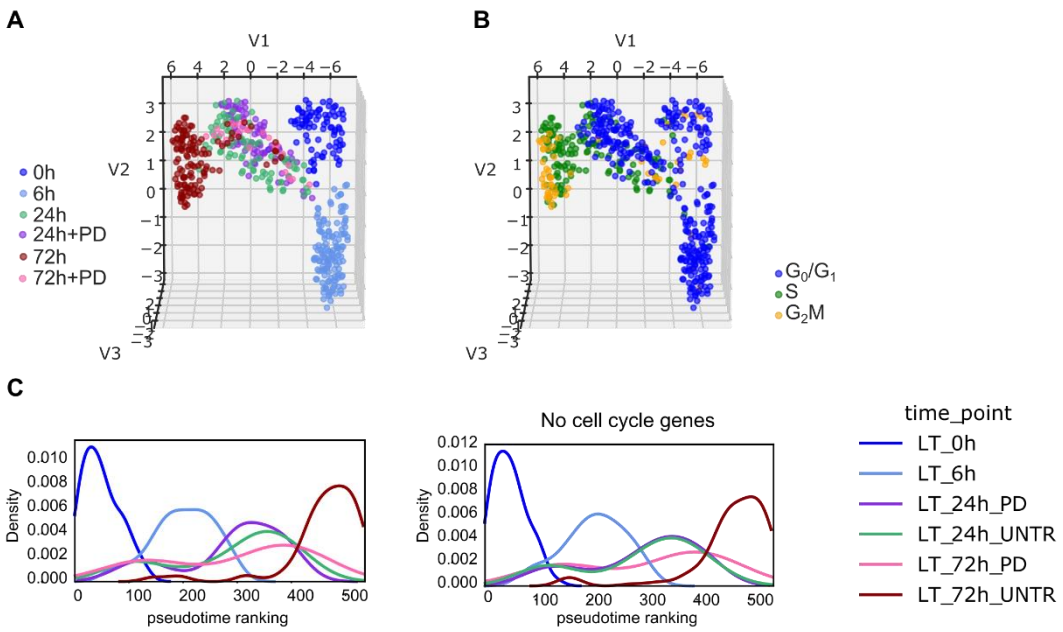


Figure 28: Analysis of time course scRNA-seq on single LT-HSCs including CDK6 inhibited LT-HSCs.

A-C) ScRNA-seq analysis of LT-HSCs cultured for 0, 6, 24 UNTR, 24 PD, 72 UNTR and 72 h PD (cells passing QC: 0 h n=85; 6 h n=134; 24 h UNTR n=86; 24 h PD n=80; 72 h UNTR n=124 cells; 72 h PD n=92). A) 3D UMAP representation of indicated populations performed on highly variable genes (14759 HVGs). Different colours indicate cells analysed at different time points. B) 3D UMAP representation as in (A) indicating the cell cycle phases assigned to each single cell. C) Left: DPT pseudotime analysis on HVGs identified between all the indicated timepoints. Cell cycle genes were included for the analysis. Right: DPT pseudotime analysis on HVGs identified between all the indicated timepoints. Cell cycle genes were excluded for the analysis.

4.11 Functional changes associated with exit from quiescence and activation are not dependent on cell cycle progression

Because previously shown transcriptomic data suggested that activation of cell growth and metabolic pathways are the first responses during exit from quiescence in LT-HSCs (Figure 29) and that the activation processes are overall similar between LT- and ST-HSCs, I decided to functionally validate these results at the cellular level in both LT-HSCs and ST-HSCs. As previously demonstrated human LT-HSCs and ST-HSCs are equally quiescent and equally small in terms of cell size and they are characterised by a low cytoplasm content, a common feature of cells residing in the G_0 quiescent state (Laurenti et al. 2015). Moreover, at the metabolic level they are characterised by low mitochondrial mass together with low mitochondrial activity, underlying the preference for a glycolytic metabolism instead of a mitochondrial oxidative metabolism (Warr and Passegué 2013; Kohli and Passegué 2014). Using the previously described *in-vitro* system (Figure 29) as the basis of these experiments I investigated the cellular effects of CDK6 inhibition after PD treatment on either LT- and ST-HSCs and I compared these to their quiescent or activated counterparts. I analysed different cellular parameters associated with stem cell quiescence and activation including cell size, mitochondrial mass and mitochondrial activity, measured as mitochondrial membrane potential before and after CDK6 inhibition (Figure 29). CB derived LT- and ST-HSCs were cultured in rich cytokine media over 72 h

with or without CDK6 inhibitor and monitored over time with live-microscopy. I could confirm that these cells are small when quiescent (0 h, ~8 µm average diameter, Figure 29 A-C), but increase significantly in size upon exit from quiescence and after 48 h in culture, with a further increase in cell size at 72 h (~15 µm average diameter, Figure 29 A-C). Moreover, this phenomenon is common to both untreated and PD treated LT-HSCs and ST-HSCs, as no significant differences between the two conditions arose from these data. In summary, because cell size increase is not affected by PD treatment in both LT- and ST-HSCs, I conclude that cell growth is independent from cell cycle entry (as defined by the phosphorylation of RB that separates early G_1 from late G_1) and the increase in HSC cell size upon activation is in agreement with what reported in the literature showing that stem cell activation is accompanied by increased cell size (van Velthoven et al. 2019).

To examine how mitochondria function changes during exit from quiescence and activation in live cells, I used two mitochondria specific dyes: the mitochondrial mass dye MitoTracker Green and TetraMethylRhodamine Methyl ester (TMRM), a cell-permeant dye that accumulates in active mitochondria with intact membrane potential. This should give a reliable idea of how mitochondria function changes during exit from quiescence and activation.

As expected, quiescent LT-HSCs and ST-HSCs showed a lower mitochondrial mass than their activated counterparts. After 24 h in culture, activated LT- and ST-HSCs had an increased mitochondrial mass, but no significant differences were detected between 24 and 48 h in culture (Figure 29 D). Similar results were obtained when measuring the mitochondrial activity with TMRM which also resulted to be increased after 24 h in culture. This increase was even higher at 48 h in both LT- and ST-HSCs, indicating that HSCs activation is accompanied by a gradual switch to mitochondria metabolism. A possible interpretation of these results is that mitochondrial mass increases upon exit from quiescence while mitochondrial function is gradually up regulated through progression into the cell cycle. Other dye-independent quantifications methods may be useful to better identify mitochondrial mass changes.

Moreover, the increase in mitochondrial mass and activity was not altered by CDK6 inhibition in both LT- and ST-HSCs, suggesting that despite blocking LT-HSCs from progressing beyond the restriction point, their metabolic switch is not linked to cell cycle progression but is instead driven by the presence of external signals.

Importantly, a recent publication reported that using dyes such as MitoTracker Green has led to inaccurate interpretations of HSCs having higher mitochondrial content than committed progenitors due to the highest capacity of HSCs to efflux these dyes (de Almeida et al. 2017). The authors however did not look at how this high mitochondrial activity of quiescent HSC changes upon activation. My data indicates alteration of mitochondrial function upon activation. Moreover, CDK6 specific inhibition had no effect on the mitochondrial function in all the analysed populations (Figure 29 E) demonstrating once again that CDK6 inhibition stops the cells in the G_0/early G_1 but leaves unaltered the upregulation of metabolic functions which is connected to the process of activation.

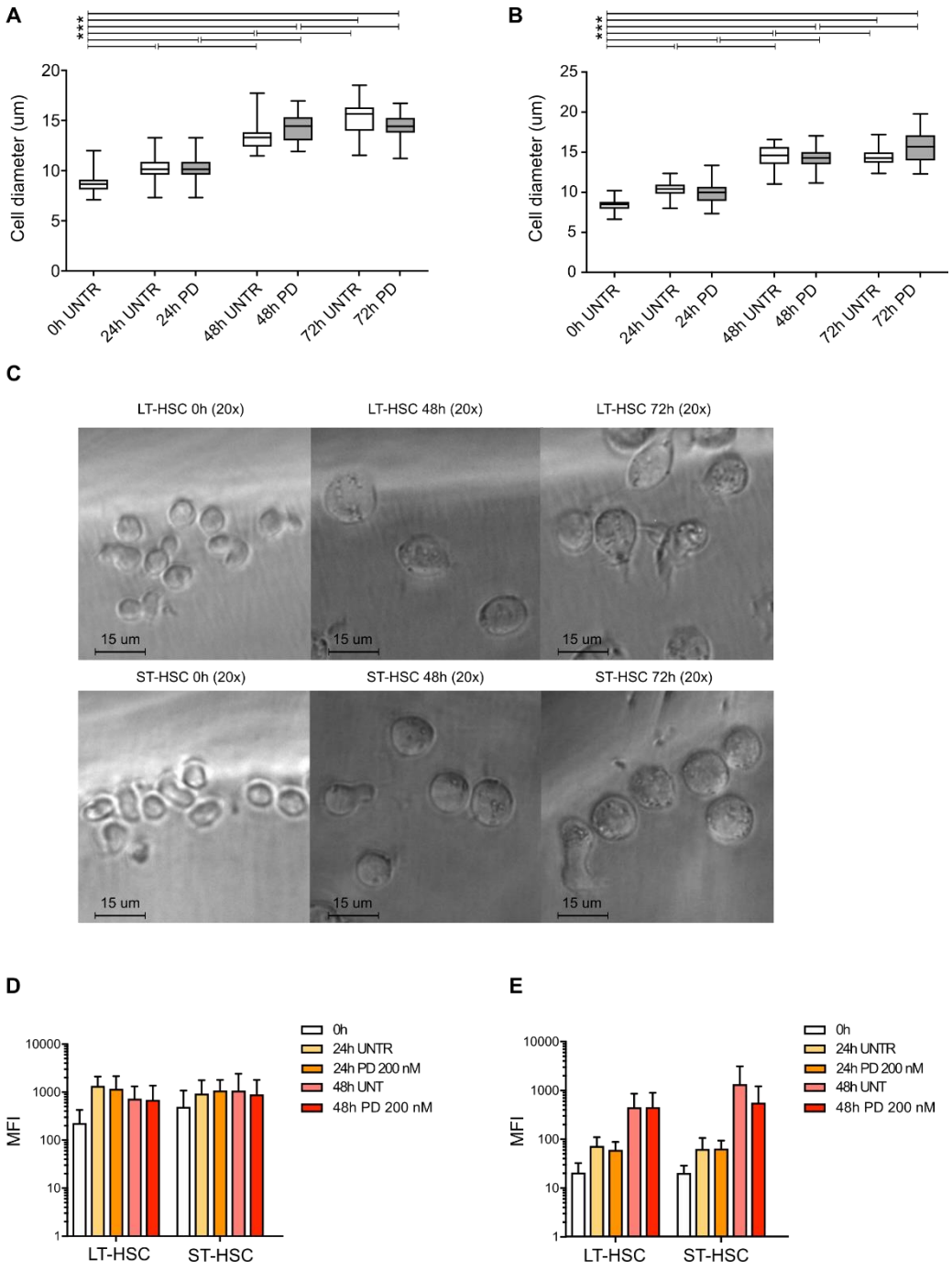


Figure 29: Changes in cell size and mitochondria metabolism accompanying HSC activation are not affected by cell cycle progression. A-B) Cell size was measured as cell diameter (µm) in LT- and ST-HSCs over time before and after PD treatment. n=2 experiments with independent CB samples. Graphs represent the cell size of n=25 cells per condition. Median, interquartile and 5-95 percentiles are shown. Statistical significance was calculated by Two-way ANOVA for multiple comparisons (***p<0.0001). No significant differences were found between untreated and PD treated conditions. C) Live-microscopy images of single LT- and ST-HSCs at 0, 48 and 72 hours in culture. Representative examples are shown. D) For each shown population MitoTracker Green Median Fluorescence Intensity (MFI) is shown on a logarithmic scale after measurement at 0, 24 and 48 h in culture with or without PD treatment (n=6 experiments with independent CB samples). Mean ± SD is shown. Statistical significance was calculated by Two-way ANOVA for multiple comparisons (ns). E) For each shown population TMRM Median Fluorescence Intensity (MFI) is shown on a logarithmic scale after measurement at 0, 24 and 48 h in culture with or without PD treatment (n=5 experiments with independent CB samples). Mean ± SD is shown. Statistical significance was calculated by Two-way ANOVA for multiple comparisons (**p<0.01). No significant differences were found between untreated and PD treated conditions.

4.12 *In-vitro* activation affects the frequency of long-term repopulating HSCs independently of cell cycle progression

Previously presented data have shown how exit of quiescence and cell cycle progression affect LT-HSCs transcriptionally and functionally. However, if these changes also affect the long-term repopulating capacity of LT-HSCs has not been investigated yet.

Based on literature and given the important changes observed in the transcriptome of LT-HSCs in correlation with their time spent in culture (Ema et al. 2000; Kent et al. 2009), differences in the repopulation capacity are expected for HSC cultured for different lengths of time, but no quantification has been reported to date at this level of resolution in the human system. Moreover, to discriminate if, in addition to time spent in culture, progression into cell cycle also affects *in-vivo* engraftment capacity or self-renewal, I investigated this using LT-HSC cultured over 24 h and 72 h with or without CDK6 inhibitor.

To test the above, I performed limiting dilutions transplantation assays (LDA) to compare the frequencies of long-term-repopulating LT-HSCs between these conditions. LT-HSCs were cultured in MEM cytokine media and three different doses of cells (700, 300 and 50 for 72 h; 300, 50 and 10 for 24 h) were transplanted intrafemorally into NSG mice. The size and composition of the grafts was measured at 18 weeks post-transplantation. Injected (IB) and non-injected (BM) bones from transplanted mice were collected, processed and analysed separately.

To assess the effects of the time spent in culture on the frequency of repopulating HSCs, I first compared the engraftment capacity (human engraftment was identified as percentage of $CD45^+$ and $GlyA^+$ cells) of untreated LT-HSCs cultured for 24 h and 72 h. Using the extreme limiting dilution analysis (ELDA) statistical method, a significant reduction in the frequency of repopulating cells ($p<0.05$) was identified between LT-HSCs cultured for 72 h compared to those cultured for 24 h. This indicates that the time spent in liquid culture in the presence of cytokines decreases the number of HSCs that can reconstitute the recipient mouse over a long time. ELDA results estimated a frequency of 1/86 and 1/217 (2.5 fold difference) long-term repopulating cells respectively within 24 h and 72 h cultured LT-HSCs (Figure 30 A-C).

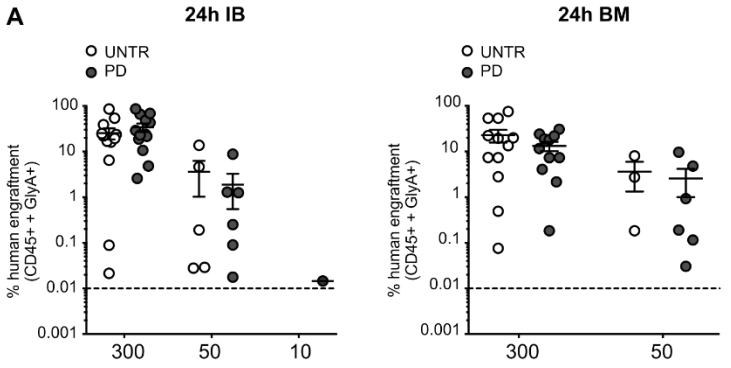
Engraftment levels in 24 h and 72 h untreated LT-HSCs were also compared to their CDK6 inhibited counterparts to investigate if the differences seen in the frequency of long-term repopulating LT-HSCs could be due the fact that these cells progress in the cell cycle while cultured *in-vitro*. As expected, engraftment levels varied based on the dose of cells injected. However, no significant differences in the percentage of human engraftment were identified between mice transplanted with untreated or PD treated LT-HSCs at 24 h at any transplanted dose (Figure 30 A). ELDA results showed no significant differences in the frequency of long-term repopulating cells between 24 h UNTR and PD treated LT-HSCs (Figure 30 D). Similarly, no differences were identified in the size of human engraftment between mice transplanted with untreated or PD treated LT-HSCs cultured for 72 h (Figure 30 B). In this case

700 and 300 doses had similar graft size distribution, indicating that 300 cells dose is probably a saturation dose for 72 h cultured LT-HSCs. LT-HSCs repopulating frequencies were also not significantly different after ELDA analysis when comparing untreated and treated conditions at 72 h (Figure 30 E). Of note, no differences in lineage composition of the grants were found in any of the conditions (data not shown). Overall, because no differences could be found in the *in-vivo* engraftment capacities of LT-HSCs that progressed into the cell cycle (untreated) and cells that were stopped by CDK6 inhibition in G_0/early G_1 *in-vitro*, these results suggest that the time spent in culture rather than cell cycle progression, affects the *in-vivo* repopulation capacity of LT-HSCs.

Moreover, while progression in the cell cycle did not affect the number of engrafting LT-HSCs, differences could still be present in the long-term self-renewal capacity of these cells.

To assess this, I performed serial transplantation assays from primary engrafted mice injected with 24 and 72 h cultured LT-HSCs. Two out of five 24 h UNTR and two out of five 24 h PD LT-HSCs secondary recipient were successfully engrafted at 12 weeks post transplantation (Figure 31 A), indicating that CDK6 inhibition over 24 h *in-vitro* doesn't confer better self-renewal properties to LT-HSCs. When looking instead at secondary recipients of 72 h untreated and 72 h PD LT-HSCs, three out of eight and one out of ten mice were respectively engrafted (Figure 31 B, not significant). Because of the small number of secondary mice engrafted, more secondary transplantation experiments will be needed to conclusively determine if blocking cell cycle progression *in-vitro* increases LT-HSC self-renewal capacity.

Overall, these results confirm that the time spent in culture, in cytokine rich media,is the major factor in determining the engraftment capacity of LT-HSCs, and not their cell cycle progression.



D

24 h			
Group	Lower	Estimate	Upper
UNTR	157	85.9	47.1
PD	114	59.8	31.2
p-value	0.435	NS	

Dose	Tested	Response	Group
300	12	12	UNTR
300	13	13	PD
50	12	5	UNTR
50	11	6	PD
10	8	0	UNTR
10	7	1	PD

E

72 h			
Group	Lower	Estimate	Upper
UNTR	352	217	134
PD	298	186	116
p-value	0.634	NS	

Dose	Tested	Response	Group
700	14	13	UNTR
700	14	12	PD
300	12	9	UNTR
300	15	14	PD
50	13	4	UNTR
50	12	5	PD

C

24 h vs 72 h			
Group	Lower	Estimate	Upper
72 h UNTR	352	217	133.6
24 h UNTR	157	86	47.1
p-value	0.0217	*	

Figure 30: Long-term repopulation capacity of LT-HSCs cultured for 24 h or 72 h *in-vitro* with or without CDK6 inhibition.
A) Percentage of human engraftment (% $CD45^{++}$ + % $GlyA^{+}$) by dose at 18 weeks after transplantation in the injected bone (IB) or bone marrow (BM) of mice injected with LT-HSCs cultured *in-vitro* for 24 h with (dark grey, n=31 mice transplanted) or without PD (white, n=31 mice transplanted). Only engrafted mice are shown. Dashed line indicated the threshold for engraftment. Mean ± SEM is shown. B) Percentage of human engraftment (%$CD45^{++}$ + % $GlyA^{+}$) by dose at 18 weeks after transplantation in the injected bone (IB) or bone marrow (BM) of mice injected with LT-HSCs cultured *in-vitro* for 72 h with (dark grey, n=42 mice transplanted) or without PD (white, n=39 mice transplanted). Only engrafted mice are shown. Dashed line indicated the threshold for engraftment. Mean ± SEM is shown. C) ELDA results table for the 24 h UNTR vs 72 h UNTR comparison. D) Top: ELDA results table for the 24 h PD vs UNTR comparison. Bottom: Doses, number of injected mice (tested), number of engrafted mice (response) and treatment groups used for ELDA analysis are shown. E) Top: ELDA results table for the 72 h PD vs UNTR comparison. Bottom: Doses, number of injected mice (tested), number of engrafted mice (response) and treatment groups used for ELDA analysis are shown.

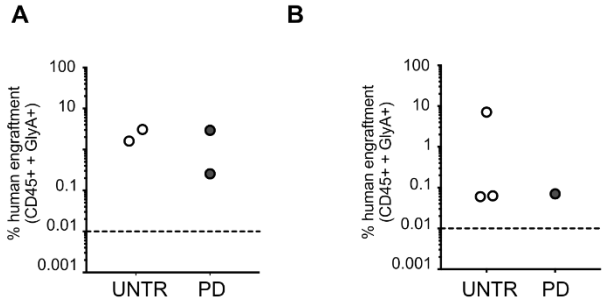


Figure 31: Long-term repopulation and self-renewal capacity of LT-HSCs cultured for 72 h *in-vitro* with or without CDK6 inhibition. A) Percentage of human engraftment (%$CD45^{++}$ + $GlyA^{+}$) in the injected bone of secondary transplantation recipients injected with $CD34^{+}$ $CD38^{-}$ cells isolated from primary mice engrafted with 24 h UNTR (white, n=5 mice transplanted) and 24 h PD treated (dark grey, n=5 mice transplanted) (A) or with 72 h UNTR (white, n=8 mice transplanted) and 72 h PD treated (dark grey, n=10 mice transplanted) (B). Only engrafted mice are shown. Dashed line indicated the threshold for engraftment.

5 DISCUSSION

The first myelo-lymphoid lineage restriction already occurs within the human haematopoietic stem cell compartment

Decades of studies on HSCs have made these cells the most characterised adult stem cells. Extensive studies in mouse (Muller-Sieburg et al. 2004; Dykstra et al. 2007; Sanjuan-Pla et al. 2013; Yamamoto et al. 2013), and more recently in human (Doulatov et al. 2010; Notta et al. 2016; Velten et al. 2017) contributed to redefine the structure of the haematopoietic hierarchy thanks to discovery of more refined and diverse HSC and progenitor populations. Heterogeneity in self-renewal properties, cycling properties and differentiation potential have been identified in both mouse and human HSCs. Indeed, more and more studies focused on investigating HSC properties at the single cell level rather than at the population level, in order to obtain a more comprehensive picture of HSC molecular and functional properties. Although there are similarities between mouse and human HSCs, species-specific differences have also been identified, highlighting the need to study human HSC directly. The purest HSCs isolated in human, to date, have been identified as $CD49f^+$ HSCs and they are characterised by the expression of Lin^- $CD34^+$ $CD38^-$ $CD45RA^-$ $CD90^+$ $CD49f^+$. However, not all these cells are capable of long-term repopulation, indicating that heterogeneity is also present within this compartment (Notta et al. 2011). In recent years, more studies are focusing on identifying new cell surface markers that would allow for better purification of purer populations of true long-term HSCs or specific HSC subsets (Knapp et al. 2018). Although the use of surface markers for HSC purification has sometimes been defines as "limiting" for the study of heterogeneous populations, it remains a very useful tool to enrich for cells with a particular behaviour. Integrating, scRNA-seq with index FACS sorting and functional assays has been proven to be the best approach for the study of the transcriptional and differentiation properties of haematopoietic cells (Paul et al. 2015; Velten et al. 2017; Karamitros et al. 2018). These approaches, in mouse, led to the identification of lineage biased HSCs, however, evidence of such populations in human has not been reported to date.

Here, I combined scRNA-seq with index FACS sorting and functional assays to comprehensively characterise the differentiation potential and transcriptome of single $CD49f^+$ HSCs. I thus combined scRNA-seq and *in-vitro* differentiation assays and I analysed more than 3000 single cells from this compartment. My findings reveal that the first lineage restriction event separates multipotent long-term repopulating cells from single cells with My/Ly but no Ery potential (Figure 18). This demonstrates that lineage restriction events already occur within the phenotypic $CD49f^+$ HSC compartment and not downstream of it as classically thought. Indirect evidence that lineage specification must occur to some degree either in the HSC or MPP compartment was already suggested by numerous recent studies that analysed progenitor cells at single cell resolution (Paul et al. 2015; Notta et al. 2016; Velten et al. 2017;

Karamitros et al. 2018). More specifically, I confirmed that a high degree of functional and molecular polarisation exists within the most immature CD49f^{+} HSC compartment along an axis of anti-correlated cell surface expression of CLEC9A and CD34 surface markers (Figure 14 and 32).

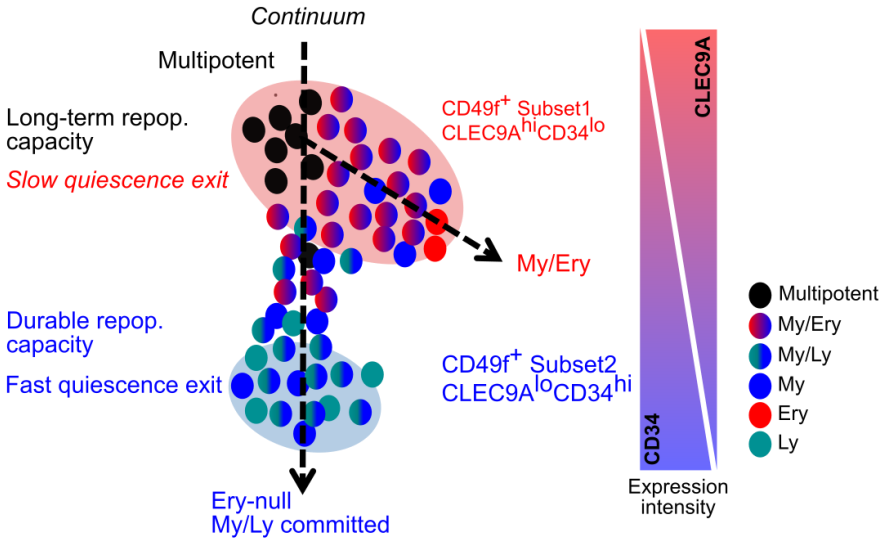


Figure 32: Graphical representation showing the proposed single-cell structure of the CD49f^{+} HSC compartment based on findings shown in this report. The figure shows a schematic representation of the continuum of differentiation and transcriptional states within the CD49f^{+} HSC compartment. The main differences identified between CD49f^{+} Subset1 and CD49f^{+} Subset2 are shown. Circles represent single cells and colours are indicative of their lineage potential. Expression intensity of CD34 and CLEC9A is shown by the arrows (blue: higher expression of CD34, red: higher expression of CLEC9A) in correlation with the single cell structure of the CD49f^{+} HSC compartment. Black arrows indicate restriction of lineage potential towards My/Ery and My/Ly lineages.

I provided evidence for a gradual developmental transition from true multipotentiality to bipotentiality occurring already within the purest HSC compartment identified to date. Thanks to the correlation of functional properties with the cell surface expression of CLEC9A and CD34, I could identify prospective purification strategies that could highly enrich for the earliest fully erythroid-deficient myelo-lymphoid restricted (Ery-null My/Ly committed) cell type reported to date (CD49f^{+} Subset2, CLEC9A^{lo} CD34^{hi}) and for the more multipotent CD49f^{+} Subset1 (CLEC9A^{hi} CD34^{lo}) population (Figure 13). Transcriptional analysis confirmed differences between the CD49f^{+} Subset2 population and LMPPs showing that the former is placed upstream of LMPPs and it thus retain properties of HSCs rather than progenitor cells (Figure 19). Additionally, the repopulation capacity of CD49f^{+} Subset2 cells extends well beyond the one of LMPPs (Goardon et al. 2011; Kohn et al. 2012; Karamitros et al. 2018). For this reason, if a name has to be given to this population, I would choose lymphoid-primed short-term repopulating HSCs.

As previously mentioned, similar data were gathered from the parallel study of the less pure HSC/MPP pool. These experiments were mainly performed by Emily Calderbank and they are not shown in this report. However, these findings were published in a co-authored manuscript (Belluschi et al. 2018) and,

together, they considerably expand the current knowledge of the molecular and cellular structure of the CD49f^{+} HSC compartment and HSC/MPP pool in humans at single cell resolution.

More in detail, the results presented in Chapter A show a comprehensive molecular and functional characterisation of the CD49f^{+} HSC compartment. Pre-existing functional and molecular heterogeneity identified in this compartment is found to be correlated with CLEC9A and CD34 cell surface expression. Anti-correlating expression of these two markers identifies a polarised but continuous structure in the differentiation output of single CD49f^{+} HSCs (Figure 11). Specifically, I identified a progression from the CLEC9A^{hi} CD34^{lo} extreme, which contains both rare cells with *in-vivo* long-term repopulating multipotent capacity and cells with *in-vitro* bilineage My/Ery potential, to the CLEC9A^{lo} CD34^{hi} extreme where single cells with *in-vitro* and *in-vivo* restricted My/Ly potential reside. Because differences in cell fate decisions were not found to be associated with the higher expression of each single surface marker, CD34^{hi} or CLEC9A^{hi} (data not shown) and because major differences in cell fate were identified only by anticorrelating expression of CLEC9A and CD34 surface markers (Figure 11), cells displaying intermediate cell surface phenotypes, such as CLEC9A^{lo} CD34^{lo} and CLEC9A^{hi} CD34^{hi}, are predicted to show heterogeneous behaviours between CD49f^{+} Subset1 and CD49f^{+} Subset2. Although, this will need to be explored further through prospective purification and *in-vitro* and *in-vivo* functional assays to confirm this hypothesis.

It is not uncommon to forget that cell surface markers used for the prospective isolation of specific cell population may also have an important function in those cells. Whereas CD34 has been used for long time as an important HSC cell surface marker, CLEC9A has been predominantly used as a dendritic cell marker (Schreibelt et al. 2012). However, little is known about the function of these two markers in HSCs. While this was beyond the scope of this study, it would still be interesting to investigate more in detail their role in CD49f^{+} HSCs ideally through specific inhibition of the expression of these two markers and subsequent *in-vitro* functional or *in-vivo* assays to assess their specific roles in HSCs. The results presented here were obtained from the study of CB derived HSCs, however, it is not to exclude that differences in the cell surface expression of these markers may be present in BM or PB derived HSCs. Preliminary data obtained by Emily Calderbank and Carys Johnson show that expression of CLECL9A is present in the BM HSC pool and CD49f^{+} HSC compartment but not in PB (data not shown). Because downregulation of cell surface markers is not unusual for circulating cells, as previously demonstrated for EPCR expressions (Cohen et al. 2015) this could also be happening for CLEC9A. Further studies are needed to investigate if CLEC9A^{hi} CD34^{lo} CD49f^{+} Subset1 and CLEC9A^{lo} CD34^{hi} CD49f^{+} Subset2 HSCs have the same functional properties as in CB, or if populations with similar properties exist are characterised by expression of different markers.

When interpreting the results obtained from single cell *in-vitro* differentiation assays presented here (Figure 10, 13, 16 and 17), it is to note that these assays may not have fully read-out the differentiation

potential of the cells analysed. These limitations need to be considered. For example, many of the single cells analysed here appeared to produce exclusively myeloid cells. A first possibility is that these cells would also differentiate into other fates (for example dendritic cells or mast cells) in other conditions or could acquire different fates under stress conditions. Evidence of direct dendritic cell specification from human HSC/MPP was recently demonstrated in a study by Lee et al. (J. Lee et al. 2017). Our laboratory has now developed single cells assays for simultaneous differentiation of single HSCs towards the myeloid and dendritic lineages or myeloid and mast cells (Emily Calderbank, unpublished). It will be of interest to determine the differentiation potential of CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells with these novel methods. Another possibility is that myeloid differentiation may simply represent a default programme when Ly- or Ery-priming is missing, as previously postulated by another publication (Iwasaki and Akashi 2007).

From the results presented here, it stands out that phenotypic HSC show many different behaviours *in-vitro*. However, this is in sharp contrast with the transcriptional similarities identified between these cells. The data presented here prove that distinct functional behaviours and restriction to bilineage differentiation potential are coordinated by an underlying transcriptional structure (Figure 15). Despite the continuity observed after dimensionality reduction analysis, differences in this structure were identified using the ICGS algorithm (Olsson et al. 2016). Three different transcriptional states could be identified within the CD49f$^+$ HSC compartment: one cluster (I) enriched in CD49f$^+$ Subset1 cells and two clusters with similar gene expression patterns, enriched in CD49f$^+$ Subset2 cells (Clusters IIa and IIb). One hypothesis is that quiescent and metabolic signatures shared by phenotypic HSC subsets (Notta et al. 2011; Laurenti et al. 2013; Chen et al. 2014) and not present in progenitor populations (Laurenti et al. 2013; Cabezas-Wallscheid et al. 2014), could be responsible of the continuous appearance of the CD49f$^+$ HSC transcriptional landscape. Data presented here demonstrate that this continuum is polarised at its extremes and changes in the transcriptional programmes of single CD49f$^+$ HSCs are associated with distinct differentiation behaviours (Figure 12 and 15). This finding highlights the importance of developing robust algorithms to identify such discrete states and has important implications for the interpretation of large single-cell transcriptomics studies.

Transcriptional differences and *in-vitro* differentiation output of CD49f$^+$ Subset1 and CD49f$^+$ Subset2 are accompanied by differences in cell cycle properties. These populations show consistently different time of first division (Figure 13). This is of interest because evidence in other stem cell types has emerged showing a correlation between the length of G_0 to G_1 transition and cell fate decisions (Pauklin and Vallier 2013; Dalton et al. 2015). Some evidence can be found in the literature (Mende et al. 2015; Grinenko et al. 2018) demonstrating such correlations in HSCs. Here, the time of first division is used as an indicative measure of the time spent in G_0/early G_1. The lengthening of the time of first division, mediated by CDK6 inhibition of CD49f$^+$ Subset2 cells has a slight but significant impact on their differentiation, showing an increase in the number of colonies including Ery and Meg cells (Figure 16).

However, this can't totally recapitulate the CD49f$^+$ Subset1 phenotype. While this may suggest that lengthening of the time spent in G_0/early G_1 increases the propensity to produce colonies of the My/Ery and My/Ery/Meg type, it is not enough to revert CD49f$^+$ Subset2 properties. Overall, these results put forward the idea that the time spent in these phases can affect, at least to some extent, HSC differentiation decisions. To confirm this, novel tools to accurately measure the duration of the G_0 to early G_1 transition will be needed. Because transcriptional data suggest that a level of lineage priming is already present within CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells it is hard to believe that the sole modulation of the cell cycle could significantly affect HSC differentiation and it is rather more plausible that, in HSCs, cell fate decisions are only slightly affected by cell cycle modulation.

To overcome the limitation in erythroid differentiation seen in the xenograft model human EPO was administered to NSG mice, unlike other recent studies on human HSC populations (Notta et al. 2011; Majeti, Park, and Weissman 2007; Knapp et al. 2018). Extending the *in-vivo* lineage read out significantly improved the understanding of the diversity of phenotypic HSC differentiation behaviours and helped with the identification of rare multipotent (Ery/My/Ly) long-term repopulating HSC within the CLEC9A^{hi} CD34^{lo} fraction of the CD49f$^+$ HSC compartment.

Similarly to all Ly-restricted or Ly-biased HSC/MPP populations previously described in mice (Dykstra et al. 2007; Cabezas-Wallscheid et al. 2014; Pietras et al. 2015; Carrelha et al. 2018) CD49f$^+$ Subset2 cells exhibit infrequent and limited repopulation capacity (Figure 18). This is in line with other studies which demonstrate that networks of TFs involved in promoting Ly-priming also inhibit HSC long term repopulation in human HSCs (Laurenti et al. 2013; Beer et al. 2014; van Galen et al. 2014). Even though the cells with long-term repopulation capacity from the CD49f$^+$ Subset2 population may not share the exact same molecular properties of all the cells within the subset, it is clear that acquisition of Ly-priming in these cells is accompanied by faster quiescence exit. In addition, the level of Ly-priming at the molecular level is lower than LMPPs (Figure 19), indicating that CD49f$^+$ Subset2 cells represent the very first step at which Ery potential is permanently lost and Ly potential specified.

In conclusion, data presented here support a model of haematopoiesis in which lymphoid lineage commitment is specified very early in quiescent and durably engrafting HSCs, as opposed to shorter lived and/or higher cycling progenitors. Lymphoid lineage is also accompanied by acquisition of a more "activated" phenotype, with shorter quiescence exit times, suggesting a transition towards a shallower or more primed quiescent state. Moreover, as differences in the niche microenvironment have been recently associated with differences in HSC bias towards My or Ly lineages in mouse (Pinho and Frenette 2019; Y. H. Ho et al. 2019), future experiment will be needed to assess if differences in the niche microenvironment and age-related niche remodelling are also responsible of the differences identified between CD49f$^+$ Subset1 and CD49f$^+$ Subset2 cells. This will have important implications in the understanding of the impact of ageing and pre-leukaemic mutations on blood formation in humans.

Cellular and molecular characterisation of exit from quiescence in human HSCs

The balance between quiescence and cell cycle entry in HSCs is finely regulated to prevent exhaustion of the HSC pool and development of haematological malignancies (Orford and Scadden 2008). Quiescence is an intrinsic property of HSCs and its regulation has been extensively studied in the mouse model (Rossi et al. 2012). However, a precise definition of HSC activation and exit from quiescence is still lacking in the field and the molecular regulation of quiescence exit has not yet been investigated in human HSCs. In the context of this study, I define exit from quiescence from the moment in which HSC are stimulated to when they phosphorylate RB (pRB) and I present a comprehensive characterisation of the molecular and cellular events occurring during activation in human HSCs, distinguishing those occurring during exit from quiescence and those requiring cell cycle progression.

In-vitro culture of HSCs results in their division. Thus, to study quiescence exit *in-vitro*, I used a model system to stop the cells in G_0/early G_1 through specific inhibition of CDK6. Division kinetics analysis of LT- and ST-HSCs recapitulated results shown in (Laurenti et al. 2015) confirming that most cells divide by 72 h in culture (Figure 21). Moreover, pRB staining assays show that exit from quiescence in these culture conditions occurs between 6 and 24 h after culture initiation in both LT- and ST-HSCs. Because CDK6 expression is higher in ST-HSCs than LT-HSCs (Laurenti et al. 2015), ST-HSC are not stopped from dividing after CDK6 inhibition and for this reason, regulation of exit from quiescence was studied mainly in LT-HSCs.

Overall, data presented here show that the transcriptional regulation of activation is very similar between LT- and ST-HSCs but more interestingly, most of the gene expression changes associated with *in-vitro* activation already occur independently of cell cycle progression (Figure 22), during exit of quiescence (before RB phosphorylation). It is known from the literature that quiescent HSCs are small in size, rely on glycolytic metabolism and have low mRNA content and activated HSCs are characterised by a switch to oxidative metabolism and active protein synthesis (Passegué et al. 2005; Simsek et al. 2010; van Velthoven et al. 2019). However, how these properties change during exit from quiescence and later cell cycle phases *in-vitro* is not known. From the data presented here, the first functions to be activated during exit from G_0 in LT-HSC are increased metabolic functions and ribosome biogenesis, together with activation of cell growth and protein synthesis, while activation of cell cycle genes is only a later response (Figure 23, 24 and 27). These results are consistent with results from studies in neural stem cells (NSCs) showing that exit from quiescence is accompanied by a remodelling of the cell metabolism (Llorens-Bobadilla et al. 2015; van Velthoven et al. 2019). Altogether, they provide first evidence that this also occurs in human HSCs independently of cell cycle progression.

Although quiescence has long been viewed as a dormant and inactive state of the cell, increasing evidence has suggested that quiescence is more like a poised state of the cell which anticipates activation, proliferation, and differentiation. This has been demonstrated in model organisms and in other stem cells to be regulated mainly at the epigenetic and post-transcriptional level (Cheung and Rando 2013; van Velthoven et al. 2019). However, experimental evidence proving that this is also valid in human HSCs is still lacking. GSEA analysis presented here, performed on bulk RNA-seq data, shows that HSC quiescence is actively maintained by the action of many TFs, most of which have been described as specific for HSC maintenance (LMO2, AHR, P53, MYB and HLF) (Cleveland et al. 2013; Bennett et al. 2015; Liu et al. 2009; Baker et al. 2014). Exit from quiescence and activation are characterised by a rapid remodelling of the activity of these TFs. This is indicative that, also in HSCs, quiescence is a poised state that allows for rapid activation. In contrast cell cycle progression in HSCs results to be driven by TF like MYC and E2F, the default TF circuity for growth in many cell types (Leone et al. 2001). Moreover, pathways involved in post-mitotic surveillance, cell cycle and mitochondria metabolism are highly enriched in activated HSCs. Metabolism related pathways only are instead enriched in LT-HSCs stopped in G_0/Early G_1 while inflammation related genes signatures (IFN, NFAT and AP1) are highly enriched in quiescent HSCs (Figure 24). These results are consistent with accumulating evidence supporting the fact that HSCs are direct targets of inflammatory signalling and that inflammatory signals may also contribute to HSC regulation in homeostatic conditions (Schuettpelz and Link 2013; Pietras 2017). Specifically, recent publications showed how the expression of IFN stimulated genes is a shared property of many stem cells, including HSCs (X. Wu et al. 2018). This was demonstrated to provide a defence mechanism for stem cells against infective agents. In accordance to data shown here, the expression of IFN stimulated genes is downregulated upon progression into the cell cycle and differentiation (Figure 27).

Bulk-RNAseq analysis provided a first characterisation of the molecular processes involved in exit from quiescence and activation in HSCs. These results were complemented with time-course transcriptomic analysis at the single cell level of LT-HSCs cultured at different time point. This is important to study exit from quiescence at the cellular level rather than at the population level, but also to discriminate if the transcriptional changes observed are related to the time spent in culture or to the fact that HSCs are blocked in the G_0 exit state (G_0/early G_1 phase). This analysis confirmed that transcriptional remodeling during quiescence exit is fairly synchronous between single cells, occurs gradually between 0 and 72 h and is mainly characterised by the up regulation of many genes, most of which are already up regulated at early time points (Figure 26). Interestingly, LT-HSCs appear to transit through a particular state, identified at 6 h in culture, before progressing into late G_1. This time point may coincide with a stress response phase where HSCs sense external signals likely associated to stress and potentially adapt to these stresses while they prepare to progress into cell cycle. 50% of the genes differentially expressed at 6 h are not differentially expressed at later time points. One hypothesis is that HSCs, after sensing

mitogenic stimuli, first enter a state similar to the G_{alert} phase, which has been described for muscle stem cells and HSCs as a priming mechanism, marked by mTORC activation, to mount a proliferative response faster in conditions of injury (Rodgers et al. 2014). It is possible that the 6 h time point corresponds to a phase where HSC prepare to exit quiescence modulating the expression of most of the genes required for G_0 exit and cell cycle progression but without actually progressing. Further experiments will be needed to assess if most cells are still in G_0 at 6 h to confirm that this time point is indicative of a different quiescence depth in LT-HSCs. In addition, analysis of mTORC activity should be carried out to formally compare the 6 h state to G_{alert}.

Gene expression pattern analysis on scRNA-seq data also shows similar findings to what has been previously identified by transcriptomic analysis performed in bulk. Genes like FOS, JUN, JUNB, CD74, GATA3, FOXP, HLX and HIF1A are among the genes identified as those rapidly down regulated during quiescence exit. As expected, metabolic, cell cycle and cell growth related pathways are gradually up regulated during activation in HSC in a cell cycle independent manner (van Velthoven et al. 2019). These findings were also confirmed functionally by mitochondrial staining and cell size measurement. Cell size, mitochondrial mass and mitochondrial function are all up regulated during exit from quiescence and progression in the cell cycle (Figure 29). Importantly, I confirmed that all these processes occur independently of cell cycle progression. Altogether, these results indicate that the metabolism response triggered by external signals and not by cell cycle progression, is an intrinsic property that drives exit from quiescence in HSCs. This is very interesting especially in the context of the role of mitochondria in HSCs and it comes in support of recent evidence demonstrating that mitochondria have an active role in regulating HSC functions and cell fate decisions independently of cell cycle progression (Vannini et al. 2016; T. T. Ho et al. 2017).

In-vitro characterisation of quiescence exit in HSCs has shown consistent remodeling of the metabolic functions of these cells. To assess if these changes affect the *in-vivo* repopulation capacity of HSCs, limiting dilution analysis was performed with LT-HSCs cultured for 24 and 72 h in NSG mice. It is still largely accepted that transition from G_0 to cell cycle leads to loss of HSC transplantation capacity (Gothot et al. 1998). My work revealed that it is the time LT-HSC spent in culture rather than their cell cycle progression that affects the frequency of long-term repopulating HSCs measured at 18 weeks from transplantation (Figure 30). This indicates that longer exposure to mitogenic signals is the main driver of the reduction of the repopulation capacity of LT-HSCs observed *in-vitro*. Because this was only demonstrated in cytokine rich media, where HSCs are forced to divide and differentiate, it is not to be excluded that HSCs cultured in different conditions may behave differently in terms of kinetics of division *in-vitro* and repopulation capacity *in-vivo*. Other culturing conditions aimed at preserving the HSC phenotype for longer times are routinely used and tested (Tajer et al. 2019) however, no *ex-vivo* culturing conditions have still been able to significantly maintain and expand human HSCs over long periods of time. Overall, because data presented here suggest that cell cycle progression is not the main

driver of differences in repopulating capacity of cultured HSCs, it can be speculated that in other culturing conditions, the repopulating capacity of HSC would still be affected depending on the time spent in culture and the type of factors present in the media, although probably to a lower extent if a reduced number of growth factors and oxygen levels are used for culturing. This could change the kinetics of activation of HSCs and reduce the pace of the metabolic activation, thus preserving better repopulation capacity *in-vivo* when compared to culture in differentiating conditions.

Moreover, *in-vivo* transplantation assays aimed at investigating the role of cell cycle progression at the base of the differences in repopulation capacity identified between LT-HSCs cultured for different times *in-vitro*, showed a slight but non-statistically significant increase in the frequency of long-term repopulating cells from untreated and CDK6 inhibited cells at both time points examined (24 h or 72 h *in-vitro*). If these results were confirmed to be significant after the analysis of a higher number of mice, CDK6 inhibition and thus a longer permanence of the cell in the early G_1 phase *in-vitro* could reduce the loss of engraftment capacity due to the time spent in culture. Useful applications could derive from the transplant of HSCs cultured *in-vitro* in the presence of CDK6 inhibitor. This would particularly benefit transplantation protocols that require prior *in-vitro* manipulation of HSCs.

In conclusion, here I provide a comprehensive characterisation of the molecular and functional changes associated with HSC activation, distinguishing events occurring during exit from quiescence from those driven by cell cycle progression. Understanding how HSC coordinate quiescence exit and cell division is of fundamental importance, as it will help to develop protocols of *ex-vivo* HSCs expansion that could be used for HSC transplants. Moreover, it may shed light on the causes of the dysregulation of the quiescence and activation balance that characterises most of the haematological malignancies. Here, results obtained from different assays all point in the same direction, indicating that metabolic regulation plays a central role in exit from quiescence and activation of HSCs. Because HSC activation occurs independently of cell cycle progression it could be hypothesised that cell cycle independent HSC activation may allow HSCs to return to quiescence or differentiate without division. However, to confirm this, further experiments will be needed. It would thus be interesting to test if this is also the case when cell cycle is inhibited through other CDK6 independent mechanisms, this could indeed be done utilising specific inhibitors of other cell cycle phases. Moreover, as CDK6 kinase independent roles have been reported (Tigan et al. 2016), it would be interesting to see if these CDK6 functions are also responsible of driving quiescence exit and are thus not detected through inhibition of CDK6 kinase activity with PD. Other inhibitors, targeting for example the TF activity of CDK6 may be used in the future to better understand the role of CDK6 in quiescence exit.

In conclusion, future studies should also focus on studying HSC metabolism more in detail as modulation of the metabolic response associated with exit from quiescence could contribute to the improvement of *ex-vivo* protocols for HSC culture.

6 BIBLIOGRAPHY

Adolfsson, Jörgen, Robert Månsson, Natalija Buza-Vidas, Anne Hultquist, Karina Liuba, Christina T. Jensen, David Bryder, et al. 2005. "Identification of Flt3 + Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential: A Revised Road Map for Adult Blood Lineage Commitment." *Cell* 121 (2): 295–306. https://doi.org/10.1016/j.cell.2005.02.013.

Ahlqvist, Kati J., Anu Suomalainen, and Riikka H. Hämäläinen. 2015. "Stem Cells, Mitochondria and Aging." *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier. https://doi.org/10.1016/j.bbabio.2015.05.014.

Akashi, Koichi, David Traver, Toshihiro Miyamoto, and Irving L. Weissman. 2000. "A Clonogenic Common Myeloid Progenitor That Gives Rise to All Myeloid Lineages." *Nature* 404 (6774): 193–97. https://doi.org/10.1038/35004599.

Almeida, Mariana Justino de, Larry L. Luchsinger, David J. Corrigan, Linda J. Williams, and Hans Willem Snoeck. 2017. "Dye-Independent Methods Reveal Elevated Mitochondrial Mass in Hematopoietic Stem Cells." *Cell Stem Cell* 21 (6): 725-729.e4. https://doi.org/10.1016/j.stem.2017.11.002.

Anders, S., P. T. Pyl, and W. Huber. 2015. "HTSeq--a Python Framework to Work with High-Throughput Sequencing Data." *Bioinformatics* 31 (2): 166–69. https://doi.org/10.1093/bioinformatics/btu638.

Anders, Simon, and Wolfgang Huber. 2010. "Differential Expression Analysis for Sequence Count Data." *Genome Biology* 11 (10): R106. https://doi.org/10.1186/gb-2010-11-10-r106.

Anjos-Afonso, Fernando, Erin Currie, Hector G Palmer, Katie E Foster, David C Taussig, and Dominique Bonnet. 2013. "CD34 - Cells at the Apex of the Human Hematopoietic Stem Cell Hierarchy Have Distinctive Cellular and Molecular Signatures." *Cell Stem Cell* 13 (2): 161–74. https://doi.org/10.1016/j.stem.2013.05.025.

Baker, Stacey J., Avi Ma'ayan, Yen K. Lieu, Premila John, M. V.Ramana Reddy, Edward Y. Chen, Qiaonan Duan, Hans Willem Snoeck, and E. Premkumar Reddy. 2014. "B-Myb Is an Essential Regulator of Hematopoietic Stem Cell and Myeloid Progenitor Cell Development." *Proceedings of the National Academy of Sciences of the United States of America* 111 (8): 3122–27. https://doi.org/10.1073/pnas.1315464111.

Becker, A.J., E.A. McCulloch, and J.E. Till. 1963. "Cytological Demonstration of the Clonal Nature of Spleen Colonies Derived from Transplanted Mouse Marrow Cells." *Nature* 197 (4866): 452–54. https://doi.org/10.1038/197452a0.

Becker, Klaus A., Prachi N. Ghule, Jaclyn A. Therrien, Jane B. Lian, Janet L. Stein, Andre J. van Wijnen, and Gary S. Stein. 2006. "Self-Renewal of Human Embryonic Stem Cells Is Supported by a Shortened G1 Cell Cycle Phase." *Journal of Cellular Physiology* 209 (3): 883–93. https://doi.org/10.1002/jcp.20776.

Beer, Philip A., David J.H.F. Knapp, Nagarajan Kannan, Paul H. Miller, Sonja Babovic, Elizabeth Bulaeva, Nima Aghaeepour, et al. 2014. "A Dominant-Negative Isoform of IKAROS Expands Primitive Normal Human Hematopoietic Cells." *Stem Cell Reports* 3 (5): 841–57. https://doi.org/10.1016/j.stemcr.2014.09.006.

Beerman, Isabel, Deepta Bhattacharya, Sasan Zandi, Mikael Sigvardsson, Irving L Weissman, David Bryder, and Derrick J Rossi. 2010. "Functionally Distinct Hematopoietic Stem Cells Modulate Hematopoietic Lineage Potential during Aging by a Mechanism of Clonal Expansion."

Proceedings of the National Academy of Sciences of the United States of America 107 (12): 5465–70. https://doi.org/10.1073/pnas.1000834107.

Belluschi, Serena, Emily F. Calderbank, Valerio Ciaurro, Blanca Pijuan-Sala, Antonella Santoro, Nicole Mende, Evangelia Diamanti, et al. 2018. "Myelo-Lymphoid Lineage Restriction Occurs in the Human Haematopoietic Stem Cell Compartment before Lymphoid-Primed Multipotent Progenitors." *Nature Communications* 9 (1): 4100. https://doi.org/10.1038/s41467-018-06442-4.

Bennett, John A., Kameshwar P. Singh, Zeenath Unnisa, Stephen L. Welle, and Thomas A. Gasiewicz. 2015. "Deficiency in Aryl Hydrocarbon Receptor (AHR) Expression throughout Aging Alters Gene Expression Profiles in Murine Long-Term Hematopoietic Stem Cells." Edited by Kevin D Bunting. *PLOS ONE* 10 (7): e0133791. https://doi.org/10.1371/journal.pone.0133791.

Benveniste, Patricia, Catherine Frelin, Salima Janmohamed, Mary Barbara, Robert Herrington, Deborah Hyam, and Norman N. Iscove. 2010. "Intermediate-Term Hematopoietic Stem Cells with Extended but Time-Limited Reconstitution Potential." *Cell Stem Cell* 6 (1): 48–58. https://doi.org/10.1016/J.STEM.2009.11.014.

Bhatia, M., J. C. Y. Wang, U. Kapp, D. Bonnet, and J. E. Dick. 1997. "Purification of Primitive Human Hematopoietic Cells Capable of Repopulating Immune-Deficient Mice." *Proceedings of the National Academy of Sciences* 94 (10): 5320–25. https://doi.org/10.1073/pnas.94.10.5320.

Bindea, Gabriela, Bernhard Mlecnik, Hubert Hackl, Pornpimol Charoentong, Marie Tosolini, Amos Kirilovsky, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski, and Jérôme Galon. 2009. "ClueGO: A Cytoscape Plug-in to Decipher Functionally Grouped Gene Ontology and Pathway Annotation Networks." *Bioinformatics* 25 (8): 1091–93. https://doi.org/10.1093/bioinformatics/btp101.

Birbrair, Alexander, and Paul S Frenette. 2016. "Niche Heterogeneity in the Bone Marrow." *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.13016.

Blomen, V. A., and J. Boonstra. 2007. "Cell Fate Determination during G1 Phase Progression." *Cellular and Molecular Life Sciences* 64 (23): 3084–3104. https://doi.org/10.1007/s00018-007-7271-z.

Boward, Ben, Tianming Wu, and Stephen Dalton. 2016. "Concise Review: Control of Cell Fate Through Cell Cycle and Pluripotency Networks." *Stem Cells* 34 (6): 1427–36. https://doi.org/10.1002/stem.2345.

Brennecke, Philip, Simon Anders, Jong Kyoung Kim, Aleksandra A. Kołodziejczyk, Xiuwei Zhang, Valentina Proserpio, Bianka Baying, et al. 2013. "Accounting for Technical Noise in Single-Cell RNA-Seq Experiments." *Nature Methods* 10 (11): 1093–98. https://doi.org/10.1038/nmeth.2645.

Bruns, Ingmar, Daniel Lucas, Sandra Pinho, Jalal Ahmed, Michele P Lambert, Yuya Kunisaki, Christoph Scheiermann, et al. 2014. "Megakaryocytes Regulate Hematopoietic Stem Cell Quiescence through CXCL4 Secretion." *Nature Medicine* 20 (11): 1315–20. https://doi.org/10.1038/nm.3707.

Busch, Katrin, Kay Klapproth, Melania Barile, Michael Flossdorf, Tim Holland-Letz, Susan M. Schlenner, Michael Reth, Thomas Höfer, and Hans-Reimer Reimer Rodewald. 2015. "Fundamental Properties of Unperturbed Haematopoiesis from Stem Cells in Vivo." *Nature* 518 (7540): 542–46. https://doi.org/10.1038/nature14242.

Cabezas-Wallscheid, Nina, Florian Buettner, Pia Sommerkamp, Daniel Klimmeck, Luisa Ladel, Frederic B. Thalheimer, Daniel Pastor-Flores, et al. 2017. "Vitamin A-Retinoic Acid Signaling

Regulates Hematopoietic Stem Cell Dormancy." *Cell* 169 (5): 807-823.e19. https://doi.org/10.1016/J.CELL.2017.04.018.

Cabezas-Wallscheid, Nina, Daniel Klimmeck, Jenny Hansson, Daniel B. B Lipka, Alejandro Reyes, Q. Qi Wang, Dieter Weichenhan, et al. 2014. "Identification of Regulatory Networks in HSCs and Their Immediate Progeny via Integrated Proteome, Transcriptome, and DNA Methylome Analysis." *Cell Stem Cell* 15 (4): 507–22. https://doi.org/10.1016/J.STEM.2014.07.005.

Calder, Ashley, Ivana Roth-Albin, Sonam Bhatia, Carlos Pilquil, Jong Hee Lee, Mick Bhatia, Marilyne Levadoux-Martin, et al. 2013. "Lengthened G1 Phase Indicates Differentiation Status in Human Embryonic Stem Cells." *Stem Cells and Development* 22 (2): 279–95. https://doi.org/10.1089/scd.2012.0168.

Calvi, L. M., G. B. Adams, K. W. Weibrecht, J. M. Weber, D. P. Olson, M. C. Knight, R. P. Martin, et al. 2003. "Osteoblastic Cells Regulate the Haematopoietic Stem Cell Niche." *Nature* 425 (6960): 841–46. https://doi.org/10.1038/nature02040.

Carrelha, Joana, Yiran Meng, Laura M. Kettyle, Tiago C. Luis, Ruggiero Norfo, Verónica Alcolea, Hanane Boukarabila, et al. 2018. "Hierarchically Related Lineage-Restricted Fates of Multipotent Haematopoietic Stem Cells." *Nature* 554 (7690): 106–11. https://doi.org/10.1038/nature25455.

Catlin, Sandra N, Lambert Busque, Rosemary E Gale, Peter Guttorp, Janis L Abkowitz, SH. Orkin, LI. Zon, et al. 2011. "The Replication Rate of Human Hematopoietic Stem Cells in Vivo." *Blood* 117 (17): 4460–66. https://doi.org/10.1182/blood-2010-08-303537.

Challen, Grant A., Nathan C. Boles, Stuart M. Chambers, and Margaret A. Goodell. 2010. "Distinct Hematopoietic Stem Cell Subtypes Are Differentially Regulated by TGF-B1." *Cell Stem Cell* 6 (3): 265–78. https://doi.org/10.1016/J.STEM.2010.02.002.

Challen, Grant A, Nathan Boles, Karen Kuan-Yin Yin Lin, and Margaret A Goodell. 2009. *Mouse Hematopoietic Stem Cell Identification and Analysis*. *Cytometry Part A*. Vol. 75. NIH Public Access. https://doi.org/10.1002/cyto.a.20674.

Chen, Lu, M. Kostadima, Joost H.A. A. Martens, Giovanni Canu, Sara P. Garcia, E. Turro, Kate Downes, et al. 2014. "Transcriptional Diversity during Lineage Commitment of Human Blood Progenitors." *Science* 345 (6204): 1251033–1251033. https://doi.org/10.1126/science.1251033.

Cheng, Jerry C., Kentaro Kinjo, Dejah R. Judelson, Jenny Chang, Winston S. Wu, Ingrid Schmid, Deepa B. Shankar, et al. 2008. "CREB Is a Critical Regulator of Normal Hematopoiesis and Leukemogenesis." *Blood* 111 (3): 1182–92. https://doi.org/10.1182/blood-2007-04-083600.

Cheung, Tom H., and Thomas A. Rando. 2013. *Molecular Regulation of Stem Cell Quiescence*. *Nature Reviews Molecular Cell Biology*. Vol. 14. Nature Publishing Group. https://doi.org/10.1038/nrm3591.

Cho, Inchul J, Prudence Pok Wai Lui, Jana Obajdin, Federica Riccio, Wladislaw Stroukov, Thea Louise Willis, Francesca Spagnoli, and Fiona M Watt. 2019. "Mechanisms, Hallmarks, and Implications of Stem Cell Quiescence." *Stem Cell Reports*. Elsevier. https://doi.org/10.1016/j.stemcr.2019.05.012.

Choe, Kevin S., Olga Ujhelly, Sandeep N. Wontakal, and Arthur I. Skoultchi. 2010. "PU.1 Directly Regulates Cdk6 Gene Expression, Linking the Cell Proliferation and Differentiation Programs in Erythroid Cells." *Journal of Biological Chemistry* 285 (5): 3044–52. https://doi.org/10.1074/jbc.M109.077727.

Civin, C I, L C Strauss, C Brovall, M J Fackler, J F Schwartz, and J H Shaper. 1984. “Antigenic Analysis of Hematopoiesis. III. A Hematopoietic Progenitor Cell Surface Antigen Defined by a Monoclonal Antibody Raised against KG-1a Cells.” *Journal of Immunology (Baltimore, Md. : 1950)* 133 (1): 157–65. http://www.ncbi.nlm.nih.gov/pubmed/6586833.

Cleveland, Susan M., Stephen Smith, Rati Tripathi, Elizabeth M. Mathias, Charnise Goodings, Natalina Elliott, Dunfa Peng, et al. 2013. “Lmo2 Induces Hematopoietic Stem Cell-Like Features in T-Cell Progenitor Cells Prior to Leukemia.” *STEM CELLS* 31 (5): 882–94. https://doi.org/10.1002/stem.1345.

Cohen, Shiri Gur, Tomer Itkin, Sagarika Chakrabarty, Claudine Graf, Orit Kollet, Aya Ludin, Karin Golan, et al. 2015. “PAR1 Signaling Regulates the Retention and Recruitment of EPCR-Expressing Bone Marrow Hematopoietic Stem Cells.” *Nature Medicine* advance on (October). https://doi.org/10.1038/nm.3960.

Copley, Michael R., Philip A. Beer, and Connie J. Eaves. 2012. “Hematopoietic Stem Cell Heterogeneity Takes Center Stage.” *Cell Stem Cell* 10 (6): 690–97. https://doi.org/10.1016/J.STEM.2012.05.006.

Copley, Michael R, and Connie J Eaves. 2013. “Developmental Changes in Hematopoietic Stem Cell Properties.” *Experimental & Molecular Medicine* 45 (11): e55–e55. https://doi.org/10.1038/emm.2013.98.

Coronado, Diana, Murielle Godet, Pierre-Yves Bourillot, Yann Tapponnier, Agnieszka Bernat, Maxime Petit, Marielle Afanassieff, et al. 2013. “A Short G1 Phase Is an Intrinsic Determinant of Naïve Embryonic Stem Cell Pluripotency.” *Stem Cell Research* 10 (1): 118–31. https://doi.org/10.1016/j.scr.2012.10.004.

Cosgun, Kadriye Nehir, Susann Rahmig, Nicole Mende, Sören Reinke, Ilona Hauber, Carola Schäfer, Anke Petzold, et al. 2014. “Kit Regulates HSC Engraftment across the Human-Mouse Species Barrier.” *Cell Stem Cell* 15 (2): 227–38. https://doi.org/10.1016/j.stem.2014.06.001.

Crisan, Mihaela, and Elaine Dzierzak. 2016. *The Many Faces of Hematopoietic Stem Cell Heterogeneity*. *Development (Cambridge)*. Vol. 143. Oxford University Press for The Company of Biologists Limited. https://doi.org/10.1242/dev.114231.

Dalton, Stephen, Y. Sela, et al., S. Pauklin, L. Vallier, Stephen Dalton, M.V. Blagosklonny, et al. 2015. “Linking the Cell Cycle to Cell Fate Decisions.” *Trends in Cell Biology* 25 (10): 592–600. https://doi.org/10.1016/j.tcb.2015.07.007.

Danet, Guénahel H., Yi Pan, Jennifer L. Luongo, Dominique A. Bonnet, and M. Celeste Simon. 2003. “Expansion of Human SCID-Repopulating Cells under Hypoxic Conditions.” *Journal of Clinical Investigation* 112 (1): 126–35.

Ding, Lei, and Sean J. Morrison. 2013. “Haematopoietic Stem Cells and Early Lymphoid Progenitors Occupy Distinct Bone Marrow Niches.” *Nature* 495 (7440): 231–35. https://doi.org/10.1038/nature11885.

Ding, Lei, Thomas L. Saunders, Grigori Enikolopov, and Sean J. Morrison. 2012. “Endothelial and Perivascular Cells Maintain Haematopoietic Stem Cells.” *Nature* 481 (7382): 457–62. https://doi.org/10.1038/nature10783.

Doulatov, Sergei, Faiyaz Notta, Kolja Eppert, Linh T Nguyen, Pamela S Ohashi, and John E Dick. 2010. “Revised Map of the Human Progenitor Hierarchy Shows the Origin of Macrophages and Dendritic

Cells in Early Lymphoid Development." *Nature Immunology* 11 (7): 585–93. https://doi.org/10.1038/ni.1889.

Doulatov, Sergei, Faiyaz Notta, Elisa Laurenti, and John E Dick. 2012. "Hematopoiesis: A Human Perspective." *Cell Stem Cell*. Elsevier Inc. https://doi.org/10.1016/j.stem.2012.01.006.

Dykstra, Brad, David Kent, Michelle Bowie, Lindsay McCaffrey, Melisa Hamilton, Kristin Lyons, Shang-Jung Lee, Ryan Brinkman, and Connie Eaves. 2007. "Long-Term Propagation of Distinct Hematopoietic Differentiation Programs In Vivo." *Cell Stem Cell* 1 (2): 218–29. https://www.sciencedirect.com/science/article/pii/S1934590907000215?via%3Dihub.

Eaves, Connie J. 2015. "Hematopoietic Stem Cells: Concepts, Definitions, and the New Reality." *Blood* 125 (17): 2605–13. https://doi.org/10.1182/blood-2014-12-570200.

Ema, Hideo, Yohei Morita, and Toshio Suda. 2014. "Heterogeneity and Hierarchy of Hematopoietic Stem Cells." *Experimental Hematology* 42 (2): 74-82.e2. https://doi.org/10.1016/j.exphem.2013.11.004.

Ema, Hideo, Hina Takano, Kazuhiro Sudo, and Hiromitsu Nakauchi. 2000. "In Vitro Self-Renewal Division of Hematopoietic Stem Cells." *Journal of Experimental Medicine* 192 (9): 1281–88. https://doi.org/10.1084/jem.192.9.1281.

Ernst, J., G. J. Nau, and Z. Bar-Joseph. 2005. "Clustering Short Time Series Gene Expression Data." *Bioinformatics* 21 (Suppl 1): i159–68. https://doi.org/10.1093/bioinformatics/bti1022.

Ernst, Jason, and Ziv Bar-Joseph. 2006. "STEM: A Tool for the Analysis of Short Time Series Gene Expression Data." *BMC Bioinformatics* 7 (1): 191. https://doi.org/10.1186/1471-2105-7-191.

Farlik, Matthias, Florian Halbritter, Fabian Müller, Fizzah A Choudry, Peter Ebert, Johanna Klughammer, Samantha Farrow, et al. 2016. "DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation." *Cell Stem Cell* 19 (6): 808–22. https://doi.org/10.1016/j.stem.2016.10.019.

Filippi, Marie-Dominique, and Saghi Ghaffari. 2019. "Mitochondria in the Maintenance of Hematopoietic Stem Cells: New Perspectives and Opportunities." *Blood* 133 (18): 1943–52. https://doi.org/10.1182/blood-2018-10-808873.

Folmes, C.D. Clifford D L C.D.L., P.P. Petras P Dzeja, T.J. Timothy J Nelson, Andre Terzic, M. Agathocleous, N.K. Love, O. Randlett, et al. 2012. "Metabolic Plasticity in Stem Cell Homeostasis and Differentiation." *Cell Stem Cell* 11 (5): 596–606. https://doi.org/10.1016/j.stem.2012.10.002.

Ford, C. E., J. L. Hamerton, D. W.H. Barnes, and J. F. Loutit. 1956. "Cytological Identification of Radiation-Chimæras." *Nature* 177 (4506): 452–54. https://doi.org/10.1038/177452a0.

Forsberg, E. Camilla, Emmanuelle Passegué, Susan S. Prohaska, Amy J. Wagers, Martina Koeva, Joshua M. Stuart, and Irving L. Weissman. 2010. "Molecular Signatures of Quiescent, Mobilized and Leukemia-Initiating Hematopoietic Stem Cells." Edited by Catherine M. Verfaillie. *PLoS ONE* 5 (1): e8785. https://doi.org/10.1371/journal.pone.0008785.

Foudi, Adlen, Konrad Hochedlinger, Denille Van Buren, Jeffrey W Schindler, Rudolf Jaenisch, Vincent Carey, and Hanno Hock. 2009. "Analysis of Histone 2B-GFP Retention Reveals Slowly Cycling Hematopoietic Stem Cells." *Nature Biotechnology* 27 (1): 84–90. https://doi.org/10.1038/nbt.1517.

Freire, Pablo R, and Orla M Conneely. 2018. “NR4A1 and NR4A3 Restrict HSC Proliferation via Reciprocal Regulation of C/EBPα and Inflammatory Signaling.” *Blood* 131 (10): 1081–93. https://doi.org/10.1182/blood-2017-07-795757.

Fujimoto, T., K. Anderson, S. E.W. Jacobsen, S. I. Nishikawa, and C. Nerlov. 2007. “Cdk6 Blocks Myeloid Differentiation by Interfering with Runx1 DNA Binding and Runx1-C/EBPα Interaction.” *EMBO Journal* 26 (9): 2361–70. https://doi.org/10.1038/sj.emboj.7601675.

Gao, Xin, Chunliang Xu, Noboru Asada, and Paul S Frenette. 2018. “The Hematopoietic Stem Cell Niche: From Embryo to Adult.” *Development* 145 (2): dev139691. https://doi.org/10.1242/dev.139691.

García-Prat, Laura, Marta Martínez-Vicente, Eusebio Perdiguero, Laura Ortet, Javier Rodríguez-Ubreva, Elena Rebollo, Vanessa Ruiz-Bonilla, et al. 2016. “Autophagy Maintains Stemness by Preventing Senescence.” *Nature* 529 (7584): 37–42. https://doi.org/10.1038/nature16187.

Georgantas, Robert W, Vivek Tanadve, Matthew Malehorn, Shelly Heimfeld, Chen Chen, Laura Carr, Francisco Martinez-Murillo, Greg Riggins, Jeanne Kowalski, and Curt I Civin. 2004. “Microarray and Serial Analysis of Gene Expression Analyses Identify Known and Novel Transcripts Overexpressed in Hematopoietic Stem Cells.” *CANCER RESEARCH*. Vol. 64. http://www.ncbi.nlm.nih.gov/GenBank/.

Goardon, Nicolas, Emanuele Marchi, Ann Atzberger, Lynn Quek, Anna Schuh, Shamit Soneji, Petter Woll, et al. 2011. “Coexistence of LMPP-like and GMP-like Leukemia Stem Cells in Acute Myeloid Leukemia.” *Cancer Cell* 19 (1): 138–52. https://www.cell.com/cancer-cell/fulltext/S1535-6108(10)00526-X.

Gonzales, Kevin Andrew Uy, Hongqing Liang, Yee Siang Lim, Yun Shen Chan, Jia Chi Yeo, Cheng Peow Tan, Bin Gao, et al. 2015. “Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways.” *Cell* 162 (3): 564–79. https://doi.org/10.1016/j.cell.2015.07.001.

Goodell, M A, K Brose, G Paradis, A S Conner, and R C Mulligan. 1996. “Isolation and Functional Properties of Murine Hematopoietic Stem Cells That Are Replicating in Vivo.” *The Journal of Experimental Medicine* 183 (4): 1797–1806. https://doi.org/10.1084/jem.183.4.1797.

Gothot, André, Johannes C.M. Van der Loo, D. Wade Clapp, and Edward F. Srour. 1998. “Cell Cycle-Related Changes in Repopulating Capacity of Human Mobilized Peripheral Blood CD34+ Cells in Non-Obese Diabetic/Severe Combined Immune- Deficient Mice.” *Blood* 92 (8): 2641–49. https://doi.org/10.1182/blood.v92.8.2641.420k36_2641_2649.

Grinenko, Tatyana, Anne Eugster, Lars Thielecke, Beáta Ramasz, Anja Krüger, Sevina Dietz, Ingmar Glauche, et al. 2018. “Hematopoietic Stem Cells Can Differentiate into Restricted Myeloid Progenitors before Cell Division in Mice.” *Nature Communications* 9 (1): 1898. https://doi.org/10.1038/s41467-018-04188-7.

Grover, Amit, Alejandra Sanjuan-Pla, Supat Thongjuea, Joana Carrelha, Alice Giustacchini, Adriana Gambardella, Iain Macaulay, et al. 2016. “Single-Cell RNA Sequencing Reveals Molecular and Functional Platelet Bias of Aged Haematopoietic Stem Cells.” *Nature Communications* 7 (1): 11075. https://doi.org/10.1038/ncomms11075.

Guo, Guoji, Sidinh Luc, Eugenio Marco, Ta-Wei Wei Lin, Cong Peng, Marc A. A. Kerenyi, Semir Beyaz, et al. 2013. “Mapping Cellular Hierarchy by Single-Cell Analysis of the Cell Surface Repertoire.” *Cell Stem Cell* 13 (4): 492–505. https://doi.org/10.1016/j.stem.2013.07.017.

Haas, Simon, Jenny Hansson, Daniel Klimmeck, Dirk Loeffler, Lars Velten, Hannah Uckelmann, Stephan Wurzer, et al. 2015. “Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors.” *Cell Stem Cell* 17 (4): 422–34. https://doi.org/10.1016/j.stem.2015.07.007.

Haas, Simon, Andreas Trumpp, and Michael D Milsom. 2018. “Causes and Consequences of Hematopoietic Stem Cell Heterogeneity.” *Cell Stem Cell* 22 (5): 627–38. https://doi.org/10.1016/j.stem.2018.04.003.

Haghverdi, Laleh, Florian Buettner, and Fabian J. Theis. 2015. “Diffusion Maps for High-Dimensional Single-Cell Analysis of Differentiation Data.” *Bioinformatics* 31 (18): 2989–98. https://doi.org/10.1093/bioinformatics/btv325.

Hamey, Fiona K., and Berthold Göttgens. 2019. “Reconstructing Gene Regulatory Networks That Control Hematopoietic Commitment.” In *Methods in Molecular Biology*, 1975:239–49. Humana, New York, NY. https://doi.org/10.1007/978-1-4939-9224-9_11.

Hamey, Fiona K, Sonia Nestorowa, Sarah J Kinston, David G Kent, Nicola K Wilson, and Berthold Göttgens. 2017. “Reconstructing Blood Stem Cell Regulatory Network Models from Single-Cell Molecular Profiles.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (23): 5822–29. https://doi.org/10.1073/pnas.1610609114.

Hao, Sha, Chen Chen, and Tao Cheng. 2016. “Cell Cycle Regulation of Hematopoietic Stem or Progenitor Cells.” *International Journal of Hematology*. Springer Japan. https://doi.org/10.1007/s12185-016-1984-4.

Hay, Stuart B., Kyle Ferchen, Kashish Chetal, H. Leighton Grimes, and Nathan Salomonis. 2018. “The Human Cell Atlas Bone Marrow Single-Cell Interactive Web Portal.” *Experimental Hematology* 68 (December): 51–61. https://doi.org/10.1016/j.exphem.2018.09.004.

Hermitte, Francis, Philippe Brunet de la Grange, Francis Belloc, Vincent Praloran, and Zoran Ivanovic. 2006. “Very Low O_2 Concentration (0.1%) Favors G_0 Return of Dividing $CD34^+$ Cells.” *Stem Cells* 24 (1): 65–73. https://doi.org/10.1634/stemcells.2004-0351.

Ho, Theodore T., Matthew R. Warr, Emmalee R. Adelman, Olivia M. Lansinger, Johanna Flach, Evgenia V. Verovskaya, Maria E. Figueroa, and Emmanuelle Passegué. 2017. “Autophagy Maintains the Metabolism and Function of Young and Old Stem Cells.” *Nature* 543 (7644): 205–10. https://doi.org/10.1038/nature21388.

Ho, Ya Hsuan, Raquel del Toro, José Rivera-Torres, Justyna Rak, Claudia Korn, Andrés García-García, David Macías, et al. 2019. “Remodeling of Bone Marrow Hematopoietic Stem Cell Niches Promotes Myeloid Cell Expansion during Premature or Physiological Aging.” *Cell Stem Cell* 25 (3): 407-418.e6. https://doi.org/10.1016/j.stem.2019.06.007.

Hock, Hanno. 2010. “Some Hematopoietic Stem Cells Are More Equal than Others.” *The Journal of Experimental Medicine* 207 (6): 1127–30. https://doi.org/10.1084/jem.20100950.

Hordyjewska, Anna, Łukasz Popiołek, and Anna Horecka. 2015. *Characteristics of Hematopoietic Stem Cells of Umbilical Cord Blood*. *Cytotechnology*. Vol. 67. Springer. https://doi.org/10.1007/s10616-014-9796-y.

Hu, Yifang, and Gordon K. Smyth. 2009. “ELDA: Extreme Limiting Dilution Analysis for Comparing Depleted and Enriched Populations in Stem Cell and Other Assays.” *Journal of Immunological Methods* 347 (1–2): 70–78. https://doi.org/10.1016/J.JIM.2009.06.008.

Ichii, Michiko, Kenji Oritani, Takafumi Yokota, Qingzhao Zhang, Karla P. Garrett, Yuzuru Kanakura, and Paul W. Kincade. 2010. "The Density of CD10 Corresponds to Commitment and Progression in the Human B Lymphoid Lineage." Edited by Derya Unutmaz. *PLoS ONE* 5 (9): e12954. https://doi.org/10.1371/journal.pone.0012954.

Itkin, Tomer, Shiri Gur-Cohen, Joel A. Spencer, Amir Schajnovitz, Saravana K. Ramasamy, Anjali P. Kusumbe, Guy Ledergor, et al. 2016. "Distinct Bone Marrow Blood Vessels Differentially Regulate Haematopoiesis." *Nature* 532 (7599): 323–28. https://doi.org/10.1038/nature17624.

Ito, Kyoko, Raphaël Turcotte, Jinhua Cui, Samuel E. Zimmerman, Sandra Pinho, Toshihide Mizoguchi, Fumio Arai, et al. 2016. "Self-Renewal of a Purified Tie2+ Hematopoietic Stem Cell Population Relies on Mitochondrial Clearance." *Science* 354 (6316): 1156–60. https://doi.org/10.1126/SCIENCE.AAF5530.

Ito, Mamoru, Hidefumi Hiramatsu, Kimio Kobayashi, Kazutomo Suzue, Mariko Kawahata, Kyoji Hioki, Yoshito Ueyama, et al. 2002. "NOD/SCID/Gamma(c)(Null) Mouse: An Excellent Recipient Mouse Model for Engraftment of Human Cells." *Blood* 100 (9): 3175–82. https://doi.org/10.1182/blood-2001-12-0207.

Ivanova, Natalia B, John T Dimos, Christoph Schaniel, Jason A Hackney, Kateri A Moore, Ihor R Lemischka, I. L. Weissman, et al. 2002. "A Stem Cell Molecular Signature." *Science (New York, N.Y.)* 298 (5593): 601–4. https://doi.org/10.1126/science.1073823.

Iwasaki, Hiromi, and Koichi Akashi. 2007. "Myeloid Lineage Commitment from the Hematopoietic Stem Cell." *Immunity*. https://doi.org/10.1016/j.immuni.2007.06.004.

Jacobson, L. O., E. L. Simmons, E. K. Marks, and J. H. Eldredge. 1951. "Recovery from Radiation Injury." *Science* 113 (2940): 510–11. https://doi.org/10.1126/science.113.2940.510.

Kamel-Reid, S, and J. Dick. 1988. "Engraftment of Immune-Deficient Mice with Human Hematopoietic Stem Cells." *Science* 242 (4886): 1706–9. https://doi.org/10.1126/science.2904703.

Karamitros, Dimitris, Bilyana Stoilova, Zahra Aboukhalil, Fiona Hamey, Andreas Reinisch, Marina Samitsch, Lynn Quek, et al. 2018. "Single-Cell Analysis Reveals the Continuum of Human Lympho-Myeloid Progenitor Cells Article." *Nature Immunology* 19 (1): 85–97. https://doi.org/10.1038/s41590-017-0001-2.

Katajisto, Pekka, Julia Döhla, Christine L Chaffer, Nalle Pentinmikko, Nemanja Marjanovic, Sharif Iqbal, Roberto Zoncu, Walter Chen, Robert A Weinberg, and David M Sabatini. 2015. "Asymmetric Apportioning of Aged Mitochondria between Daughter Cells Is Required for Stemness." *Science* 348 (6232): 340–43. https://doi.org/10.1126/science.1260384.

Kataoka, Keisuke, Tomohiko Sato, Akihide Yoshimi, Susumu Goyama, Takako Tsuruta, Hiroshi Kobayashi, Munetake Shimabe, et al. 2011. "Evi1 Is Essential for Hematopoietic Stem Cell Self-Renewal, and Its Expression Marks Hematopoietic Cells with Long-Term Multilineage Repopulating Activity." *The Journal of Experimental Medicine* 208 (12): 2403–16. https://doi.org/10.1084/jem.20110447.

Kawamoto, H, K Ohmura, S Fujimoto, and Y Katsura. 1999. "Emergence of T Cell Progenitors without B Cell or Myeloid Differentiation Potential at the Earliest Stage of Hematopoiesis in the Murine Fetal Liver." *Journal of Immunology (Baltimore, Md. : 1950)* 162 (5): 2725–31. http://www.ncbi.nlm.nih.gov/pubmed/10072517.

Kawamoto, Hiroshi, Tomokatsu Ikawa, Kyoko Masuda, Haruka Wada, and Yoshimoto Katsura. 2010. "A Map for Lineage Restriction of Progenitors during Hematopoiesis: The Essence of the Myeloid-

Based Model." *Immunological Reviews* 238 (1): 23–36. https://doi.org/10.1111/j.1600-065X.2010.00959.x.

Kent, David G, Michael R Copley, Claudia Benz, Stefan Wöhrer, Brad J Dykstra, Elaine Ma, John Cheyne, et al. 2009. "Prospective Isolation and Molecular Characterization of Hematopoietic Stem Cells with Durable Self-Renewal Potential." *Blood* 113 (25): 6342–50. https://doi.org/10.1182/blood-2008-12-192054.

Kheir, Tony Bou, and Anders H. Lund. 2010. "Epigenetic Dynamics across the Cell Cycle." *Essays in Biochemistry* 48 (September): 107–20. https://doi.org/10.1042/BSE0480107.

Knapp, David J.H.F. H. F., Colin A. Hammond, Tony Hui, Marijn T.J. J. van Loenhout, Fangwu Wang, Nima Aghaeepour, Paul H. Miller, et al. 2018. "Single-Cell Analysis Identifies a CD33 + Subset of Human Cord Blood Cells with High Regenerative Potential." *Nature Cell Biology* 20 (6): 710–20. https://doi.org/10.1038/s41556-018-0104-5.

Kobayashi, Chiharu I., and Toshio Suda. 2012. "Regulation of Reactive Oxygen Species in Stem Cells and Cancer Stem Cells." *Journal of Cellular Physiology* 227 (2): 421–30. https://doi.org/10.1002/jcp.22764.

Kohli, Latika, and Emmanuelle Passegué. 2014. "Surviving Change: The Metabolic Journey of Hematopoietic Stem Cells." *Trends in Cell Biology* 24 (8): 479–87. https://doi.org/10.1016/j.tcb.2014.04.001.

Kohn, Lisa A, Qian-Lin Hao, Rajkumar Sasidharan, Chintan Parekh, Shundi Ge, Yuhua Zhu, Hanna K A Mikkola, and Gay M Crooks. 2012. "Lymphoid Priming in Human Bone Marrow Begins before Expression of CD10 with Upregulation of L-Selectin." *Nature Immunology* 13 (10): 963–71. https://doi.org/10.1038/ni.2405.

Kollmann, Karoline, Gerwin Heller, Christine Schneckenleithner, Wolfgang Warsch, Ruth Scheicher, Rene G. Ott, Markus Schäfer, et al. 2013. "A Kinase-Independent Function of CDK6 Links the Cell Cycle to Tumor Angiogenesis." *Cancer Cell* 24 (2): 167–81. https://doi.org/10.1016/j.ccr.2013.07.012.

Komorowska, Karolina, Alexander Doyle, Martin Wahlestedt, Agatheeswaran Subramaniam, Shubhranshu Debnath, Jun Chen, Shamit Soneji, et al. 2017. "Hepatic Leukemia Factor Maintains Quiescence of Hematopoietic Stem Cells and Protects the Stem Cell Pool during Regeneration." *Cell Reports* 21 (12): 3514–23. https://doi.org/10.1016/j.celrep.2017.11.084.

Kondo, Motonari, Irving L. Weissman, and Koichi Akashi. 1997. "Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow." *Cell* 91 (5): 661–72. https://doi.org/10.1016/S0092-8674(00)80453-5.

Kozar, Katarzyna, and Piotr Sicinski. 2005. "Cell Cycle Progression without Cyclin D-CDK4 and Cyclin D-CDK6 Complexes." *Cell Cycle*. Taylor and Francis Inc. https://doi.org/10.4161/cc.4.3.1551.

Lange, Christian, and Federico Calegari. 2010. "Cdks and Cyclins Link G1 Length and Differentiation of Embryonic, Neural and Hematopoietic Stem Cells." *Cell Cycle*. Taylor & Francis. https://doi.org/10.4161/cc.9.10.11598.

Lansdorp, P M, H J Sutherland, and C J Eaves. 1990. "Selective Expression of CD45 Isoforms on Functional Subpopulations of CD34+ Hemopoietic Cells from Human Bone Marrow." *The Journal of Experimental Medicine* 172 (1): 363–66. https://doi.org/10.1084/jem.172.1.363.

Larochelle, A., M. Savona, M. Wiggins, S. Anderson, B. Ichwan, K. Keyvanfar, S. J. Morrison, and C. E. Dunbar. 2011. “Human and Rhesus Macaque Hematopoietic Stem Cells Cannot Be Purified Based Only on SLAM Family Markers.” *Blood* 117 (5): 1550–54. https://doi.org/10.1182/blood-2009-03-212803.

Laurenti, Elisa, and John E Dick. 2012. “Molecular and Functional Characterization of Early Human Hematopoiesis.” *Annals of the New York Academy of Sciences* 1266 (1): 68–71. https://doi.org/10.1111/j.1749-6632.2012.06577.x.

Laurenti, Elisa, Sergei Doulatov, Sasan Zandi, Ian Plumb, Jing Chen, Craig April, Jian-Bing Fan, and John E. Dick. 2013. “The Transcriptional Architecture of Early Human Hematopoiesis Identifies Multilevel Control of Lymphoid Commitment.” *Nature Immunology* 14 (7): 756–63. https://doi.org/10.1038/ni.2615.

Laurenti, Elisa, Catherine Frelin, Stephanie Xie, Robin Ferrari, Cyrille F. F. Dunant, Sasan Zandi, Andrea Neumann, et al. 2015. “CDK6 Levels Regulate Quiescence Exit in Human Hematopoietic Stem Cells.” *Cell Stem Cell* 16 (3): 302–13. https://doi.org/10.1016/j.stem.2015.01.017.

Laurenti, Elisa, and Berthold Göttgens. 2018. “From Haematopoietic Stem Cells to Complex Differentiation Landscapes.” *Nature*. Nature Publishing Group. https://doi.org/10.1038/nature25022.

Laurenti, Elisa, Barbara Varnum-Finney, Anne Wilson, Isabel Ferrero, William E. Blanco-Bose, Armin Ehninger, Paul S. Knoepfler, et al. 2008. “Hematopoietic Stem Cell Function and Survival Depend on C-Myc and N-Myc Activity.” *Cell Stem Cell* 3 (6): 611–24. https://doi.org/10.1016/j.stem.2008.09.005.

Lee, Jaeyop, Yu Jerry Zhou, Wenji Ma, Wanwei Zhang, Arafat Aljoufi, Thomas Luh, Kimberly Lucero, et al. 2017. “Lineage Specification of Human Dendritic Cells Is Marked by IRF8 Expression in Hematopoietic Stem Cells and Multipotent Progenitors.” *Nature Immunology* 18 (8): 877–88. https://doi.org/10.1038/ni.3789.

Lee, Sung-Young, Jaeho Yoon, Mee-Hyun Lee, Sung Keun Jung, Dong Joon Kim, Ann M Bode, Jaebong Kim, and Zigang Dong. 2012. “The Role of Heterodimeric AP-1 Protein Comprised of JunD and c-Fos Proteins in Hematopoiesis.” *The Journal of Biological Chemistry* 287 (37): 31342–48. https://doi.org/10.1074/jbc.M112.387266.

Leone, Gustavo, Rosalie Sears, Erich Huang, Rachel Rempel, Faison Nuckolls, Chi Hyun Park, Paloma Giangrande, et al. 2001. “Myc Requires Distinct E2F Activities to Induce S Phase and Apoptosis.” *Molecular Cell* 8 (1): 105–13. https://doi.org/10.1016/S1097-2765(01)00275-1.

Li, Rudong, Yin Wang, Hui Cheng, Gang Liu, Tao Cheng, Yunlong Liu, and Lei Liu. 2017. “System Modeling Reveals the Molecular Mechanisms of HSC Cell Cycle Alteration Mediated by Maff and Egr3 under Leukemia.” *BMC Systems Biology* 11 (S5): 91. https://doi.org/10.1186/s12918-017-0467-4.

Lim, Shuhui, and Philipp Kaldis. 2013. “Cdks, Cyclins and CKIs: Roles beyond Cell Cycle Regulation.” *Development* 140 (15): 3079–93. https://doi.org/10.1242/dev.091744.

Liu, Yan, Shannon E. Elf, Yasuhiko Miyata, Goro Sashida, Yuhui Liu, Gang Huang, Silvana Di Giandomenico, et al. 2009a. “P53 Regulates Hematopoietic Stem Cell Quiescence.” *Cell Stem Cell* 4 (1): 37–48. https://doi.org/10.1016/J.STEM.2008.11.006.

———. 2009b. “P53 Regulates Hematopoietic Stem Cell Quiescence.” *Cell Stem Cell* 4 (1): 37–48. https://doi.org/10.1016/j.stem.2008.11.006.

Llorens-Bobadilla, Enric, Sheng Zhao, Avni Baser, Gonzalo Saiz-Castro, Klara Zwadlo, and Ana Martin-Villalba. 2015. “Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells That Become Activated upon Brain Injury.” *Cell Stem Cell* 17 (3): 329–40. https://doi.org/10.1016/j.stem.2015.07.002.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12). https://doi.org/10.1186/s13059-014-0550-8.

Luchsinger, Larry L., Mariana Justino de Almeida, David J. Corrigan, Melanie Mumau, and Hans-Willem Snoeck. 2016. “Mitofusin 2 Maintains Haematopoietic Stem Cells with Extensive Lymphoid Potential.” *Nature* 529 (7587): 528–31. https://doi.org/10.1038/nature16500.

Luchsinger, Larry L., Alexandros Strikoudis, Nichole M. Danzl, Erin C. Bush, Michael O. Finlayson, Prakash Satwani, Megan Sykes, Masayuki Yazawa, and Hans-Willem Snoeck. 2019. “Harnessing Hematopoietic Stem Cell Low Intracellular Calcium Improves Their Maintenance In Vitro.” *Cell Stem Cell*, June. https://doi.org/10.1016/j.stem.2019.05.002.

Ludin, Aya, Shiri Gur-Cohen, Karin Golan, Kerstin B. Kaufmann, Tomer Itkin, Chiara Medaglia, Xin Jiang Lu, Guy Ledergor, Orit Kollet, and Tsvee Lapidot. 2014. “Reactive Oxygen Species Regulate Hematopoietic Stem Cell Self-Renewal, Migration and Development, as Well as Their Bone Marrow Microenvironment.” *Antioxidants and Redox Signaling*. Mary Ann Liebert Inc. https://doi.org/10.1089/ars.2014.5941.

Majeti, Ravindra, Christopher Y. Park, and Irving L. Weissman. 2007. “Identification of a Hierarchy of Multipotent Hematopoietic Progenitors in Human Cord Blood.” *Cell Stem Cell* 1 (6): 635–45. https://doi.org/10.1016/j.stem.2007.10.001.

Malumbres, Marcos, and Mariano Barbacid. 2009. “Cell Cycle, CDKs and Cancer: A Changing Paradigm.” *Nature Reviews Cancer*. Nature Publishing Group. https://doi.org/10.1038/nrc2602.

Malumbres, Marcos, Rocío Sotillo, David Santamaría, Javier Galán, Ana Cerezo, Sagrario Ortega, Pierre Dubus, and Mariano Barbacid. 2004. “Mammalian Cells Cycle without the D-Type Cyclin-Dependent Kinases Cdk4 and Cdk6.” *Cell* 118 (4): 493–504. https://doi.org/10.1016/j.cell.2004.08.002.

Månsson, Robert, Anne Hultquist, Sidinh Luc, Liping Yang, Kristina Anderson, Shabnam Kharazi, Suleiman Al-Hashmi, et al. 2007. “Molecular Evidence for Hierarchical Transcriptional Lineage Priming in Fetal and Adult Stem Cells and Multipotent Progenitors.” *Immunity* 26 (4): 407–19. https://doi.org/10.1016/j.immuni.2007.02.013.

Matsumoto, Akinobu, Shoichiro Takeishi, Tomoharu Kanie, Etsuo Susaki, Ichiro Onoyama, Yuki Tateishi, Keiko Keiichi I. Nakayama, and Keiko Keiichi I. Nakayama. 2011. “P57 Is Required for Quiescence and Maintenance of Adult Hematopoietic Stem Cells.” *Cell Stem Cell* 9 (3): 262–71. https://www.sciencedirect.com/science/article/pii/S1934590911003298?via%3Dihub.

Mende, Nicole, Erika E. Kuchen, Mathias Lesche, Tatyana Grinenko, Konstantinos D. Kokkaliaris, Helmut Hanenberg, Dirk Lindemann, et al. 2015. “CCND1–CDK4–Mediated Cell Cycle Progression Provides a Competitive Advantage for Human Hematopoietic Stem Cells in Vivo.” *The Journal of Experimental Medicine* 212 (8): 1171–83. https://doi.org/10.1084/jem.20150308.

Méndez-Ferrer, Simón, Tatyana V. Michurina, Francesca Ferraro, Amin R. Mazloom, Ben D. MacArthur, Sergio A. Lira, David T. Scadden, Avi Ma’ayan, Grigori N. Enikolopov, and Paul S. Frenette. 2010. “Mesenchymal and Haematopoietic Stem Cells Form a Unique Bone Marrow Niche.” *Nature* 466 (7308): 829–34. https://doi.org/10.1038/nature09262.

Miller, Cindy L., Brad Dykstra, and Connie J. Eaves. 2008. “Characterization of Mouse Hematopoietic Stem and Progenitor Cells.” In *Current Protocols in Immunology*, 80:22B.2.1-22B.2.31. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/0471142735.im22b02s80.

Moignard, Victoria, and Berthold Göttgens. 2014. “Transcriptional Mechanisms of Cell Fate Decisions Revealed by Single Cell Expression Profiling.” *BioEssays* 36 (4): 419–26. https://doi.org/10.1002/bies.201300102.

Moignard, Victoria, Iain C. MacAulay, Gemma Swiers, Florian Buettner, Judith Schütte, Fernando J. Calero-Nieto, Sarah Kinston, et al. 2013. “Characterization of Transcriptional Networks in Blood Stem and Progenitor Cells Using High-Throughput Single-Cell Gene Expression Analysis.” *Nature Cell Biology* 15 (4): 363–72. https://doi.org/10.1038/ncb2709.

Moore, M. A. S., N. Williams, and D. Metcalf. 1973. “In Vitro Colony Formation by Normal and Leukemic Human Hematopoietic Cells: Characterization of the Colony-Forming Cells 2.” *JNCI: Journal of the National Cancer Institute* 50 (3): 603–23. https://doi.org/10.1093/jnci/50.3.603.

Morita, Yohei, Hideo Ema, and Hiromitsu Nakauchi. 2010. “Heterogeneity and Hierarchy within the Most Primitive Hematopoietic Stem Cell Compartment.” *The Journal of Experimental Medicine* 207 (6): 1173–82. https://doi.org/10.1084/jem.20091318.

Mortensen, Monika, Elizabeth J. Soilleux, Gordana Djordjevic, Rebecca Tripp, Michael Lutteropp, Elham Sadighi-Akha, Amanda J. Stranks, et al. 2011. “The Autophagy Protein Atg7 Is Essential for Hematopoietic Stem Cell Maintenance.” *The Journal of Experimental Medicine* 208 (3): 455–67. https://doi.org/10.1084/jem.20101145.

Muller-Sieburg, Christa E, Rebecca H Cho, Lars Karlsson, Jing-F Huang, and Hans B Sieburg. 2004. “Myeloid-Biased Hematopoietic Stem Cells Have Extensive Self-Renewal Capacity but Generate Diminished Lymphoid Progeny with Impaired IL-7 Responsiveness.” *Blood* 103 (11): 4111–18. https://doi.org/10.1182/blood-2003-10-3448.

Müller-Sieburg, Christa E, Rebecca H Cho, Marilyn Thoman, Becky Adkins, and Hans B Sieburg. 2002. “Deterministic Regulation of Hematopoietic Stem Cell Self-Renewal and Differentiation.” *Blood* 100 (4): 1302–9. http://www.ncbi.nlm.nih.gov/pubmed/12149211.

Muller-Sieburg, Christa E, Hans B Sieburg, Jeff M Bernitz, and Giulio Cattarossi. 2012. “Stem Cell Heterogeneity: Implications for Aging and Regenerative Medicine.” *Blood* 119 (17): 3900–3907. https://doi.org/10.1182/blood-2011-12-376749.

Nakamura-Ishizu, Ayako, Hitoshi Takizawa, and Toshio Suda. 2014. “The Analysis, Roles and Regulation of Quiescence in Hematopoietic Stem Cells.” *Development (Cambridge, England)* 141 (24): 4656–66. https://doi.org/10.1242/dev.106575.

Nakamura-Ishizu, Ayako, Keiyo Takubo, Masato Fujioka, and Toshio Suda. 2014. “Megakaryocytes Are Essential for HSC Quiescence through the Production of Thrombopoietin.” *Biochemical and Biophysical Research Communications* 454 (2): 353–57. https://doi.org/10.1016/J.BBRC.2014.10.095.

Neganova, I, X Zhang, S Atkinson, and M Lako. 2009. “Expression and Functional Analysis of G1 to S Regulatory Components Reveals an Important Role for CDK2 in Cell Cycle Regulation in Human Embryonic Stem Cells.” *Oncogene* 28 (1): 20–30. https://doi.org/10.1038/onc.2008.358.

Nestorowa, Sonia, Fiona K Hamey, Blanca Pijuan Sala, Evangelia Diamanti, Mairi Shepherd, Elisa Laurenti, Nicola K Wilson, David G Kent, and Berthold Göttgens. 2016. “A Single-Cell Resolution

Map of Mouse Hematopoietic Stem and Progenitor Cell Differentiation." *Blood* 128 (8): e20–31. https://doi.org/10.1182/blood-2016-05-716480.

Ng, Ashley P, and Warren S Alexander. 2017. "Haematopoietic Stem Cells: Past, Present and Future." *Cell Death Discovery*. Nature Publishing Group. https://doi.org/10.1038/cddiscovery.2017.2.

Notta, Faiyaz, Sergei Doulatov, Elisa Laurenti, Armando Poeppl, Igor Jurisica, John E. Dick, M. Sauvageau, et al. 2011. "Isolation of Single Human Hematopoietic Stem Cells Capable of Long-Term Multilineage Engraftment." *Science* 333 (6039): 218–21. https://doi.org/10.1126/science.1201219.

Notta, Faiyaz, Sasan Zandi, Naoya Takayama, Stephanie Dobson, Olga I. Gan, Gavin Wilson, Kerstin B. Kaufmann, et al. 2016. "Distinct Routes of Lineage Development Reshape the Human Blood Hierarchy across Ontogeny." *Science* 351 (6269): aab2116. https://doi.org/10.1126/science.aab2116.

Novershtern, Noa, Aravind Subramanian, Lee N. Lawton, Raymond H. Mak, W. Nicholas Haining, Marie E. McConkey, Naomi Habib, et al. 2011. "Densely Interconnected Transcriptional Circuits Control Cell States in Human Hematopoiesis." *Cell* 144 (2): 296–309. https://doi.org/10.1016/j.cell.2011.01.004.

Oguro, Hideyuki, Atsushi Iwama, Yohei Morita, Takehiko Kamijo, Maarten van Lohuizen, and Hiromitsu Nakauchi. 2006. "Differential Impact of Ink4a and Arf on Hematopoietic Stem Cells and Their Bone Marrow Microenvironment in Bmi1-Deficient Mice." *The Journal of Experimental Medicine* 203 (10): 2247–53. https://doi.org/10.1084/jem.20052477.

Olsson, Andre, Meenakshi Venkatasubramanian, Viren K. Chaudhri, Bruce J. Aronow, Nathan Salomonis, Harinder Singh, and H. Leighton Grimes. 2016. "Single-Cell Analysis of Mixed-Lineage States Leading to a Binary Cell Fate Choice." *Nature* 537 (7622): 698–702. https://doi.org/10.1038/nature19348.

Omatsu, Yoshiki, Tatsuki Sugiyama, Hiroshi Kohara, Gen Kondoh, Nobutaka Fujii, Kenji Kohno, and Takashi Nagasawa. 2010. "The Essential Functions of Adipo-Osteogenic Progenitors as the Hematopoietic Stem and Progenitor Cell Niche." *Immunity* 33 (3): 387–99. https://doi.org/10.1016/J.IMMUNI.2010.08.017.

Orford, Keith W., and David T. Scadden. 2008. "Deconstructing Stem Cell Self-Renewal: Genetic Insights into Cell-Cycle Regulation." *Nat Rev Genet* 9 (2): 115–28. https://doi.org/10.1038/nrg2269.

Orkin, Stuart H, and Leonard I Zon. 2008. "Hematopoiesis: An Evolving Paradigm for Stem Cell Biology." *Cell*. Howard Hughes Medical Institute. https://doi.org/10.1016/j.cell.2008.01.025.

Pantano, Lorena. 2019. "DEGreport: Report of DEG Analysis. R Package Version 1.20.0. Http://Lpantano.Github.Io/DEGreport/."

Park, Jong-Eun, Krzysztof Polański, Kerstin Meyer, and Sarah A. Teichmann. 2018. "Fast Batch Alignment of Single Cell Transcriptomes Unifies Multiple Mouse Cell Atlases into an Integrated Landscape." *BioRxiv*, 397042. https://doi.org/10.1101/397042.

Passegué, Emmanuelle, Amy J. Wagers, Sylvie Giuriato, Wade C. Anderson, and Irving L. Weissman. 2005. "Global Analysis of Proliferation and Cell Cycle Gene Expression in the Regulation of Hematopoietic Stem and Progenitor Cell Fates." *Journal of Experimental Medicine* 202 (11): 1599–1611. https://doi.org/10.1084/jem.20050967.

Pauklin, Siim, and Ludovic Vallier. 2013. “The Cell-Cycle State of Stem Cells Determines Cell Fate Propensity.” *Cell* 155 (6): 1338. https://doi.org/10.1016/j.cell.2013.08.031.

Paul, Franziska, Ya’ara Arkin, Amir Giladi, Diego Adhemar Adhemar Jaitin, Ephraim Kenigsberg, Hadas Keren-Shaul, Deborah Winter, et al. 2015. “Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors.” *Cell* 163 (7): 1663–77. https://doi.org/10.1016/j.cell.2015.11.013.

Pellin, Danilo, Mariana Loperfido, Cristina Baricordi, Samuel L. Wolock, Annita Montepeloso, Olga K. Weinberg, Alessandra Biffi, Allon M. Klein, and Luca Biasco. 2019. “A Comprehensive Single Cell Transcriptional Landscape of Human Hematopoietic Progenitors.” *Nature Communications* 10 (1). https://doi.org/10.1038/s41467-019-10291-0.

Perié, Leïla, Ken R. Duffy, Lianne Kok, Rob J. De Boer, and Ton N. Schumacher. 2015. “The Branching Point in Erythro-Myeloid Differentiation.” *Cell* 163 (7): 1655–62. https://doi.org/10.1016/j.cell.2015.11.059.

Picelli, Simone, Åsa K. Björklund, Omid R. Faridani, Sven Sagasser, Gösta Winberg, and Rickard Sandberg. 2013. “Smart-Seq2 for Sensitive Full-Length Transcriptome Profiling in Single Cells.” *Nature Methods* 10 (11): 1096–98. https://doi.org/10.1038/nmeth.2639.

Pietras, Eric M. 2017. “Inflammation: A Key Regulator of Hematopoietic Stem Cell Fate in Health and Disease.” *Blood* 130 (15): 1693–98. https://doi.org/10.1182/blood-2017-06-780882.

Pietras, Eric M., Damien Reynaud, Yoon-A A. Kang, Daniel Carlin, Fernando J. Calero-Nieto, Andrew D. Leavitt, Joshua A. M Stuart, et al. 2015. “Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions.” *Cell Stem Cell* 17 (1): 35–46. https://doi.org/10.1016/j.stem.2015.05.003.

Pietras, Eric M., Matthew R. Warr, and Emmanuelle Passegué. 2011. “Cell Cycle Regulation in Hematopoietic Stem Cells.” *The Journal of Cell Biology* 195 (5): 709–20. https://doi.org/10.1083/jcb.201102131.

Pina, Cristina, Cristina Fugazza, Alex J. Tipping, John Brown, Shamit Soneji, Jose Teles, Carsten Peterson, and Tariq Enver. 2012. “Inferring Rules of Lineage Commitment in Haematopoiesis.” *Nature Cell Biology* 14 (3): 287–94. https://doi.org/10.1038/ncb2442.

Pina, Cristina, Gillian May, Shamit Soneji, Dengli Hong, and Tariq Enver. 2008. “MLLT3 Regulates Early Human Erythroid and Megakaryocytic Cell Fate.” *Cell Stem Cell* 2 (3): 264–73. https://doi.org/10.1016/J.STEM.2008.01.013.

Pinho, Sandra, and Paul S. Frenette. 2019. “Haematopoietic Stem Cell Activity and Interactions with the Niche.” *Nature Reviews Molecular Cell Biology* 20 (5): 303–20. https://doi.org/10.1038/s41580-019-0103-9.

Pinho, Sandra, Julie Lacombe, Maher Hanoun, Toshihide Mizoguchi, Ingmar Bruns, Yuya Kunisaki, and Paul S Frenette. 2013. “PDGFRα and CD51 Mark Human Nestin+ Sphere-Forming Mesenchymal Stem Cells Capable of Hematopoietic Progenitor Cell Expansion.” *The Journal of Experimental Medicine* 210 (7): 1351–67. https://doi.org/10.1084/jem.20122252.

Pinho, Sandra, Tony Marchand, Eva Yang, Qiaozhi Wei, Claus Nerlov, and Paul S Frenette Correspondence. 2018. “Lineage-Biased Hematopoietic Stem Cells Are Regulated by Distinct Niches.” *Developmental Cell* 44: 634–41. https://doi.org/10.1016/j.devcel.2018.01.016.

Pluznik, Dov H., and Leo Sachs. 2005. “The Cloning of Normal ‘Mast’ Cells in Tissue Culture.” *Journal of Cellular and Comparative Physiology* 66 (3): 319–24. https://doi.org/10.1002/jcp.1030660309.

Rodgers, Joseph T., Katherine Y. King, Jamie O. Brett, Melinda J. Cromie, Gregory W. Charville, Katie K. Maguire, Christopher Brunson, et al. 2014. “MTORC1 Controls the Adaptive Transition of Quiescent Stem Cells from G0 to GAlert.” *Nature* 510 (7505): 393–96. https://doi.org/10.1038/nature13255.

Rodriguez-Fraticelli, Alejo E., Samuel L. Wolock, Caleb S. Weinreb, Riccardo Panero, Sachin H. Patel, Maja Jankovic, Jianlong Sun, Raffaele A. Calogero, Allon M. Klein, and Fernando D. Camargo. 2018. “Clonal Analysis of Lineage Fate in Native Haematopoiesis.” *Nature* 553 (7687): 212–16. https://doi.org/10.1038/nature25168.

Rossi, Lara, Kuanyin K. K. Lin, Nathan C. C. Boles, Liubin Yang, Katherine Y. Y. King, Mira Jeong, Allison Mayle, and Margaret A. A. Goodell. 2012. “Less Is More: Unveiling the Functional Core of Hematopoietic Stem Cells through Knockout Mice.” *Cell Stem Cell* 11 (3): 302–17. https://doi.org/10.1016/j.stem.2012.08.006.

Rubin, Seth M. 2012. “Deciphering the Rb Phosphorylation Code.” https://doi.org/10.1016/j.tibs.2012.10.007.

Ruiz, Sergio, Athanasia D. Panopoulos, Aída Herrerías, Karl-Dimiter Bissig, Margaret Lutz, W. Travis Berggren, Inder M. Verma, and Juan Carlos Izpisua Belmonte. 2011. “A High Proliferation Rate Is Required for Cell Reprogramming and Maintenance of Human Embryonic Stem Cell Identity.” *Current Biology* 21 (1): 45–52. https://doi.org/10.1016/J.CUB.2010.11.049.

Sakaue-Sawano, Asako, Hiroshi Kurokawa, Toshifumi Morimura, Aki Hanyu, Hiroshi Hama, Hatsuki Osawa, Saori Kashiwagi, et al. 2008. “Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression.” *Cell* 132 (3): 487–98. https://doi.org/10.1016/j.cell.2007.12.033.

Sanjuan-Pla, Alejandra, Iain C. Macaulay, Christina T. Jensen, Petter S. Woll, Tiago C. Luis, Adam Mead, Susan Moore, et al. 2013. “Platelet-Biased Stem Cells Reside at the Apex of the Haematopoietic Stem-Cell Hierarchy.” *Nature* 502 (7470): 232–36. https://doi.org/10.1038/nature12495.

Satija, Rahul, Jeffrey A. Farrell, David Gennert, Alexander F. Schier, and Aviv Regev. 2015. “Spatial Reconstruction of Single-Cell Gene Expression Data.” *Nature Biotechnology* 33 (5): 495–502. https://doi.org/10.1038/nbt.3192.

Schreibelt, Gerty, Lieke J.J. Klinkenberg, Luis J. Cruz, Paul J. Tacken, Jurjen Tel, Martin Kreutz, Gosse J. Adema, Gordon D. Brown, Carl G. Figdor, and I. Jolanda M. De Vries. 2012. “The C-Type Lectin Receptor CLEC9A Mediates Antigen Uptake and (Cross-)Presentation by Human Blood BDCA3+ Myeloid Dendritic Cells.” *Blood* 119 (10): 2284–92. https://doi.org/10.1182/blood-2011-08-373944.

Schuettpelz, Laura G., and Daniel C. Link. 2013. “Regulation of Hematopoietic Stem Cell Activity by Inflammation.” *Frontiers in Immunology* 4. https://doi.org/10.3389/fimmu.2013.00204.

Seita, Jun, and Irving L. Weissman. 2010. *Hematopoietic Stem Cell: Self-Renewal versus Differentiation. Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. Vol. 2. http://doi.wiley.com/10.1002/wsbm.86.

Semenza, Gregg L. 2011. “Hypoxia-Inducible Factor 1: Regulator of Mitochondrial Metabolism and Mediator of Ischemic Preconditioning.” *Biochimica et Biophysica Acta* 1813 (7): 1263–68. https://doi.org/10.1016/j.bbamcr.2010.08.006.

Shima, Haruko, Keiyo Takubo, Naoko Tago, Hiroko Iwasaki, Fumio Arai, Takao Takahashi, and Toshio Suda. 2010. "Acquisition of G_0 State by CD34-Positive Cord Blood Cells after Bone Marrow Transplantation." *Experimental Hematology* 38 (12): 1231–40. https://doi.org/10.1016/j.exphem.2010.08.004.

Simsek, Tugba, Fatih Kocabas, Junke Zheng, Ralph J. R.J. Ralph J DeBerardinis, Ahmed I. Mahmoud, Eric N. Olson, Jay W. Schneider, et al. 2010. "The Distinct Metabolic Profile of Hematopoietic Stem Cells Reflects Their Location in a Hypoxic Niche." *Cell Stem Cell* 7 (3): 380–90. https://doi.org/10.1016/j.stem.2010.07.011.

Singh, Amar M, James Chappell, Robert Trost, Li Lin, Tao Wang, Jie Tang, Brittany K Matlock, et al. 2013. "Cell-Cycle Control of Developmentally Regulated Transcription Factors Accounts for Heterogeneity in Human Pluripotent Cells." *Stem Cell Reports* 1 (6): 532–44. https://doi.org/10.1016/j.stemcr.2013.10.009.

Sitnicka, Ewa, Natalija Buza-Vidas, Staffan Larsson, Jens M Nygren, Karina Liuba, and Sten Erik W Jacobsen. 2003. "Human CD34+ Hematopoietic Stem Cells Capable of Multilineage Engrafting NOD/SCID Mice Express Flt3: Distinct Flt3 and c-Kit Expression and Response Patterns on Mouse and Candidate Human Hematopoietic Stem Cells." *Blood* 102 (3): 881–86. https://doi.org/10.1182/blood-2002-06-1694.

Snoeck, Hans Willem. 2017. "Mitochondrial Regulation of Hematopoietic Stem Cells." *Current Opinion in Cell Biology*. Elsevier Current Trends. https://doi.org/10.1016/j.ceb.2017.12.010.

Spangrude, G J, S Heimfeld, and I L Weissman. 1988. "Purification and Characterization of Mouse Hematopoietic Stem Cells." *Science (New York, N.Y.)* 241 (4861): 58–62. https://doi.org/10.1126/science.2898810.

Srivastava, Surabhi, Rakesh K. Mishra, and Jyotsna Dhawan. 2010. "Regulation of Cellular Chromatin State: Insights from Quiescence and Differentiation." *Organogenesis*. Taylor & Francis. https://doi.org/10.4161/org.6.1.11337.

Subramanian, Aravind, Pablo Tamayo, Vamsi K. Mootha, Sayan Mukherjee, Benjamin L. Ebert, Michael A. Gillette, Amanda Paulovich, et al. 2005. "Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles." *Proceedings of the National Academy of Sciences of the United States of America* 102 (43): 15545–50. https://doi.org/10.1073/pnas.0506580102.

Suda, Toshio, Keiyo Takubo, and Gregg L. Semenza. 2011. *Metabolic Regulation of Hematopoietic Stem Cells in the Hypoxic Niche*. *Cell Stem Cell*. Vol. 9. Cell Press. https://doi.org/10.1016/j.stem.2011.09.010.

Sun, Jianlong, Azucena Ramos, Brad Chapman, Jonathan B. Johnnidis, Linda Le, Yu-Jui Ho, Allon Klein, Oliver Hofmann, and Fernando D. Camargo. 2014. "Clonal Dynamics of Native Haematopoiesis." *Nature* 514 (7522): 322–27. https://doi.org/10.1038/nature13824.

Sutherland, H J, C J Eaves, P M Lansdorp, J D Thacker, and D E Hogge. 1991. "Differential Regulation of Primitive Human Hematopoietic Cells in Long-Term Cultures Maintained on Genetically Engineered Murine Stromal Cells." *Blood* 78 (3): 666–72. http://www.ncbi.nlm.nih.gov/pubmed/1713512.

Szilvassy, S J, R K Humphries, P M Lansdorp, A C Eaves, and C J Eaves. 1990. "Quantitative Assay for Totipotent Reconstituting Hematopoietic Stem Cells by a Competitive Repopulation Strategy." *Proceedings of the National Academy of Sciences of the United States of America* 87 (22). http://www.ncbi.nlm.nih.gov/pubmed/2247442.

Tajer, Parisa, Karin Pike-Overzet, Sagrario Arias, Menzo Havenga, and Frank Staal. 2019. “Ex Vivo Expansion of Hematopoietic Stem Cells for Therapeutic Purposes: Lessons from Development and the Niche.” *Cells* 8 (2): 169. https://doi.org/10.3390/cells8020169.

Takubo, Keiyo, Nobuhito Goda, Wakako Yamada, Hirono Iriuchishima, Eiji Ikeda, Yoshiaki Kubota, Haruko Shima, et al. 2010. “Regulation of the HIF-1α Level Is Essential for Hematopoietic Stem Cells.” *Cell Stem Cell* 7 (3): 391–402. https://doi.org/10.1016/j.stem.2010.06.020.

Terzi, Menderes Yusuf, Muzeyyen Izmirli, and Bulent Gogebakan. 2016. “The Cell Fate: Senescence or Quiescence.” *Molecular Biology Reports*. Springer Netherlands. https://doi.org/10.1007/s11033-016-4065-0.

Tigan, A. S., F. Bellutti, K. Kollmann, G. Tebb, and V. Sexl. 2016. “CDK6-a Review of the Past and a Glimpse into the Future: From Cell-Cycle Control to Transcriptional Regulation.” *Oncogene*. Nature Publishing Group. https://doi.org/10.1038/onc.2015.407.

Till, J.E. E, E.A. A McCulloch, and L. Siminovitch. 1964. “A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 2 (1): 683. https://doi.org/10.1073/pnas.51.1.29.

Tirosh, Itay, Benjamin Izar, Sanjay M. Prakadan, Marc H. Wadsworth, Daniel Treacy, John J. Trombetta, Asaf Rotem, et al. 2016. “Dissecting the Multicellular Ecosystem of Metastatic Melanoma by Single-Cell RNA-Seq.” *Science* 352 (6282): 189–96. https://doi.org/10.1126/science.aad0501.

Tothova, Zuzana, Ramya Kollipara, Brian J. Huntly, Benjamin H. Lee, Diego H. Castrillon, Dana E. Cullen, Elizabeth P. McDowell, et al. 2007. “FoxOs Are Critical Mediators of Hematopoietic Stem Cell Resistance to Physiologic Oxidative Stress.” *Cell* 128 (2): 325–39. https://doi.org/10.1016/J.CELL.2007.01.003.

Uchida, Naoyuki, Brad Dykstra, KJ. Kristin J Lyons, Frank Y.K FY. Leung, and CJ. Connie J Eaves. 2003. “Different in Vivo Repopulating Activities of Purified Hematopoietic Stem Cells before and after Being Stimulated to Divide in Vitro with the Same Kinetics” 31 (12): 1338–47. https://doi.org/10.1016/j.exphem.2003.09.001.

Umemoto, Terumasa, Michihiro Hashimoto, Takayoshi Matsumura, Ayako Nakamura-Ishizu, and Toshio Suda. 2018. “Ca 2+-Mitochondria Axis Drives Cell Division in Hematopoietic Stem Cells.” *J. Exp. Med* 215 (8): 2097–2113. https://doi.org/10.1084/jem.20180421.

Unnisa, Zeenath, Jason P. Clark, Jayeeta Roychoudhury, Elizabeth Thomas, Lino Tessarollo, Neal G. Copeland, Nancy A. Jenkins, H. Leighton Grimes, and Ashish R. Kumar. 2012. “Meis1 Preserves Hematopoietic Stem Cells in Mice by Limiting Oxidative Stress.” *Blood* 120 (25): 4973–81. https://doi.org/10.1182/blood-2012-06-435800.

Vallier, Ludovic. 2015. “Cell Cycle Rules Pluripotency.” *Cell Stem Cell*. https://doi.org/10.1016/j.stem.2015.07.019.

van Galen, Peter, Antonija Kreso, Erno Wienholds, Elisa Laurenti, Kolja Eppert, Eric R. Lechman, Nathan Mbong, et al. 2014. “Reduced Lymphoid Lineage Priming Promotes Human Hematopoietic Stem Cell Expansion.” *Cell Stem Cell* 14 (1): 94–106. https://doi.org/10.1016/J.STEM.2013.11.021.

Vannini, Nicola, Vasco Campos, Mukul Girotra, Vincent Trachsel, Shanti Rojas-Sutterlin, Josefine Tratwal, Simone Ragusa, et al. 2019. “The NAD-Booster Nicotinamide Riboside Potently

Stimulates Hematopoiesis through Increased Mitochondrial Clearance." *Cell Stem Cell* 24 (3): 405-418.e7. https://doi.org/10.1016/J.STEM.2019.02.012.

Vannini, Nicola, Mukul Girotra, Olaia Naveiras, Gennady Nikitin, Vasco Campos, Sonja Giger, Aline Roch, Johan Auwerx, and Matthias P. Lutolf. 2016. "Specification of Haematopoietic Stem Cell Fate via Modulation of Mitochondrial Activity." *Nature Communications* 7 (1): 13125. https://doi.org/10.1038/ncomms13125.

Velten, Lars, Simon F. Haas, Simon Raffel, Sandra Blaszkiewicz, Saiful Islam, Bianca P. Hennig, Christoph Hirche, et al. 2017. "Human Haematopoietic Stem Cell Lineage Commitment Is a Continuous Process." *Nature Cell Biology* 19 (4): 271–81. https://doi.org/10.1038/ncb3493.

Velthoven, Cindy T.J. van, Antoine de Morree, Ingrid M. Egner, Jamie O. Brett, and Thomas A. Rando. 2017. "Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo." *Cell Reports* 21 (7): 1994–2004. https://doi.org/10.1016/J.CELREP.2017.10.037.

Velthoven, Cindy T J van, Thomas A. Rando, Cindy T.J. van Velthoven, and Thomas A. Rando. 2019. "Stem Cell Quiescence: Dynamism, Restraint, and Cellular Idling Cindy" 24 (2): 213–25. https://www.cell.com/cell-stem-cell/fulltext/S1934-5909(19)30001-3?dgcid=raven_jbs_etoc_email.

Viatour, Patrick, Tim C. Somervaille, Shivkumar Venkatasubrahmanyam, Scott Kogan, Margaret E. McLaughlin, Irving L. Weissman, Atul J. Butte, Emmanuelle Passegué, and Julien Sage. 2008. "Hematopoietic Stem Cell Quiescence Is Maintained by Compound Contributions of the Retinoblastoma Gene Family." *Cell Stem Cell* 3 (4): 416–28. https://doi.org/10.1016/j.stem.2008.07.009.

Wang, J., T. Kimura, R. Asada, S. Harada, S. Yokota, Y. Kawamoto, Y. Fujimura, T. Tsuji, S. Ikehara, and Y. Sonoda. 2003. "SCID-Repopulating Cell Activity of Human Cord Blood-Derived CD34- Cells Assured by Intra-Bone Marrow Injection." *Blood* 101 (8): 2924–31. https://doi.org/10.1182/blood-2002-09-2782.

Wang, J C, M Doedens, and J E Dick. 1997. "Primitive Human Hematopoietic Cells Are Enriched in Cord Blood Compared with Adult Bone Marrow or Mobilized Peripheral Blood as Measured by the Quantitative in Vivo SCID-Repopulating Cell Assay." *Blood* 89 (11): 3919–24. http://www.ncbi.nlm.nih.gov/pubmed/9166828.

Warr, Matthew R., and Emmanuelle Passegué. 2013. "Metabolic Makeover for HSCs." *Cell Stem Cell* 12 (1): 1–3. https://doi.org/10.1016/j.stem.2012.12.005.

Wilkinson, Adam C., and Berthold Göttgens. 2013. "Transcriptional Regulation of Haematopoietic Stem Cells." In , 187–212. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6621-1_11.

Wilson, Anne, Elisa Laurenti, Gabriela Oser, Richard C. van der Wath, William Blanco-Bose, Maike Jaworski, Sandra Offner, et al. 2008. "Hematopoietic Stem Cells Reversibly Switch from Dormancy to Self-Renewal during Homeostasis and Repair." *Cell* 135 (6): 1118–29. https://doi.org/10.1016/j.cell.2008.10.048.

Wilson, Anne, Mark J. Murphy, Thordur Oskarsson, Konstantinos Kaloulis, Michael D. Bettess, Gabriela M. Oser, Anne Catherine Pasche, Christian Knabenhans, H. Robson MacDonald, and Andreas Trumpp. 2004. "C-Myc Controls the Balance between Hematopoietic Stem Cell Self-Renewal and Differentiation." *Genes and Development* 18 (22): 2747–63. https://doi.org/10.1101/gad.313104.

Wilson, Nicola K., David G. Kent, Florian Buettner, Mona Shehata, Iain C. Macaulay, Fernando J. Calero-Nieto, Manuel S??nchez Castillo, et al. 2015. “Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations.” *Cell Stem Cell* 16 (6): 712–24. https://doi.org/10.1016/j.stem.2015.04.004.

Wolf, F. Alexander, Philipp Angerer, and Fabian J. Theis. 2018. “SCANPY: Large-Scale Single-Cell Gene Expression Data Analysis.” *Genome Biology* 19 (1): 15. https://doi.org/10.1186/s13059-017-1382-0.

Wu, T. D., and S. Nacu. 2010. “Fast and SNP-Tolerant Detection of Complex Variants and Splicing in Short Reads.” *Bioinformatics* 26 (7): 873–81. https://doi.org/10.1093/bioinformatics/btq057.

Wu, Xianfang, Viet Loan Dao Thi, Yumin Huang, Eva Billerbeck, Debjani Saha, Hans Heinrich Hoffmann, Yaomei Wang, et al. 2018. “Intrinsic Immunity Shapes Viral Resistance of Stem Cells.” *Cell* 172 (3): 423-438.e25. https://doi.org/10.1016/j.cell.2017.11.018.

Yamamoto, Ryo, Yohei Morita, Jun Ooehara, Sanae Hamanaka, Masafumi Onodera, Karl Lenhard Rudolph, Hideo Ema, and Hiromitsu Nakauchi. 2013. “Clonal Analysis Unveils Self-Renewing Lineage-Restricted Progenitors Generated Directly from Hematopoietic Stem Cells.” *Cell* 154 (5): 1112–26. https://doi.org/10.1016/J.CELL.2013.08.007.

Yu, Wen Mei, Xia Liu, Jinhua Shen, Olga Jovanovic, Elena E. Pohl, Stanton L. Gerson, Toren Finkel, Hal E. Broxmeyer, and Cheng Kui Qu. 2013. “Metabolic Regulation by the Mitochondrial Phosphatase PTPMT1 Is Required for Hematopoietic Stem Cell Differentiation.” *Cell Stem Cell* 12 (1): 62–74. https://doi.org/10.1016/j.stem.2012.11.022.

Yuan, Youzhong, Hongmei Shen, David S. Franklin, David T. Scadden, and Tao Cheng. 2004. “In Vivo Self-Renewing Divisions of Haematopoietic Stem Cells Are Increased in the Absence of the Early G1-Phase Inhibitor, P18INK4C.” *Nature Cell Biology* 6 (5): 436–42. https://doi.org/10.1038/ncb1126.

Zerbino, Daniel R., Premanand Achuthan, Wasiu Akanni, M. Ridwan Amode, Daniel Barrell, Jyothish Bhai, Konstantinos Billis, et al. 2018. “Ensembl 2018.” *Nucleic Acids Research* 46 (D1): D754–61. https://doi.org/10.1093/nar/gkx1098.

Zhang, Hongbo, Keir J. Menzies, and Johan Auwerx. 2018. “The Role of Mitochondria in Stem Cell Fate and Aging.” *Development* 145 (8): dev143420. https://doi.org/10.1242/dev.143420.

Zhang, Jiwang, Chao Niu, Ling Ye, Haiyang Huang, Xi He, Wei-Gang Tong, Jason Ross, et al. 2003. “Identification of the Haematopoietic Stem Cell Niche and Control of the Niche Size.” *Nature* 425 (6960): 836–41. https://doi.org/10.1038/nature02041.

Zhao, Meng, and Linheng Li. 2016. “Dissecting the Bone Marrow HSC Niches.” *Cell Research* 26 (9): 975–76. https://doi.org/10.1038/cr.2016.71.

Zhao, Meng, John M Perry, Heather Marshall, Aparna Venkatraman, Pengxu Qian, Xi C He, Jasimuddin Ahamed, and Linheng Li. 2014. “Megakaryocytes Maintain Homeostatic Quiescence and Promote Post-Injury Regeneration of Hematopoietic Stem Cells.” *Nature Medicine* 20 (11): 1321–26. https://doi.org/10.1038/nm.3706.

Zheng, Shiwei, Efthymia Papalexi, Andrew Butler, William Stephenson, and Rahul Satija. 2018. “Molecular Transitions in Early Progenitors during Human Cord Blood Hematopoiesis.” *Molecular Systems Biology* 14 (3). https://doi.org/10.15252/msb.20178041.

Zou, Jing, Ping Zou, Jie Wang, Lei Li, Yong Wang, Daohong Zhou, and Lingbo Liu. 2012. “Inhibition of P38 MAPK Activity Promotes Ex Vivo Expansion of Human Cord Blood Hematopoietic Stem Cells.” *Annals of Hematology* 91 (6): 813–23. https://doi.org/10.1007/s00277-011-1397-7.

Zou, Peng, Hiroki Yoshihara, Kentaro Hosokawa, Ikue Tai, Kaori Shinmyozu, Fujiko Tsukahara, Yoshiro Maru, Keiko Keiichi I. Nakayama, Keiko Keiichi I. Nakayama, and Toshio Suda. 2011. “P57Kip2 and P27Kip1 Cooperate to Maintain Hematopoietic Stem Cell Quiescence through Interactions with Hsc70.” *Cell Stem Cell* 9 (3): 247–61. https://www.sciencedirect.com/science/article/pii/S193459091100333X?via%3Dihub.